(12) UK Patent Application (19) GB (11) 2 318 021 (13) A

(43) Date of A Publication 08.04.1998

(21) Application No 9720970.4

(22) Date of Filing 02.10.1997

(30) Priority Data
(31) 08726329 (32) 03.10.1996 (33) US

(71) Applicant(s)
Motorola Inc

(Incorporated in USA - Delaware)

1303 East Algonquin Road, Schaumburg, Illinois 60196, United States of America

(72) Inventor(s)
Timothy John Burke
Douglas J Newlin

(74) Agent and/or Address for Service
Sarah Gibson
Motorola Limited, European Intellectual Property Operation, Midpoint, Alencon Link, BASINGSTOKE, Hampshire, RG21 7PL, United Kingdom

(51) INT CL^6
H04N 7/14, H04M 3/56

(52) UK CL (Edition P)
H4K KF56 KOD3

(56) Documents Cited
GB 2294377 A GB 2280822 A WO 95/12281 A1
US 5541640 A US 5524110 A US 5473366 A

(58) Field of Search
UK CL (Edition O) H4K KF56 KF56A KOD3
INT CL^6 H04M 3/56 11/08, H04N 7/14 7/15

(54) **Wireline audio and video conferencing and telephony**

(57) A video access apparatus (110, 150) provides for audio and video teleconferencing and telephony via a first communication channel (103) having communication with a network (140), such as the public switched telephone network or an ISDN network. The video access apparatus (110) includes a wireline network interface (210); a radio frequency modulator/demodulator (205); a user interface (215); and a processor arrangement (190). A videophone apparatus (700, 800) is coupleable to a video access apparatus via a second communication channel for video reception and transmission, and via a third communication channel for audio reception and transmission. The videophone apparatus includes a video monitor (715), a camera interface (235), a video camera (720), and a telephony module (710). Multiple videophone apparatuses (700, 800) may be used simultaneously, and multiple video signals from the videophone apparatuses (700, 800) may be multiplexed and combined into one composite video signal for transmission to the network (140).

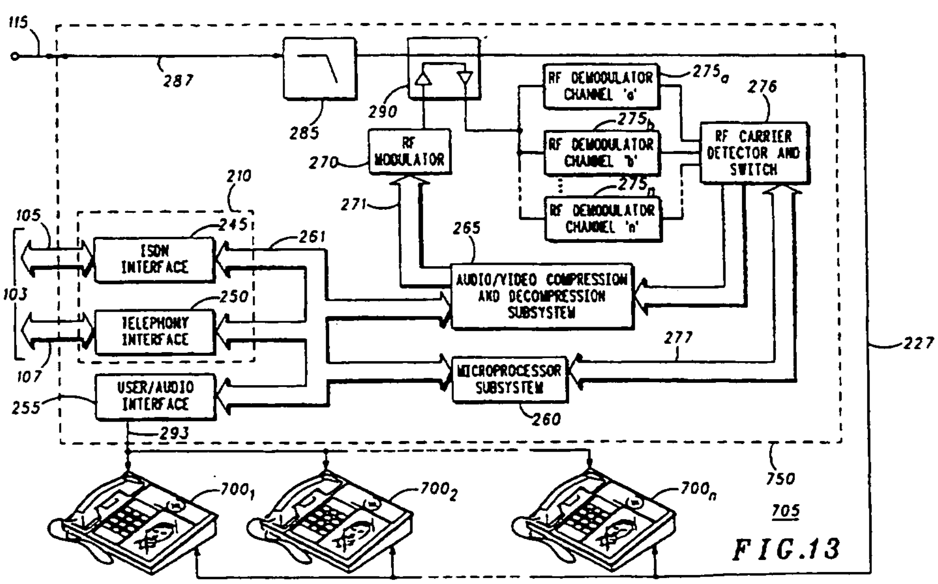


FIG.13

GB 2 318 021 A

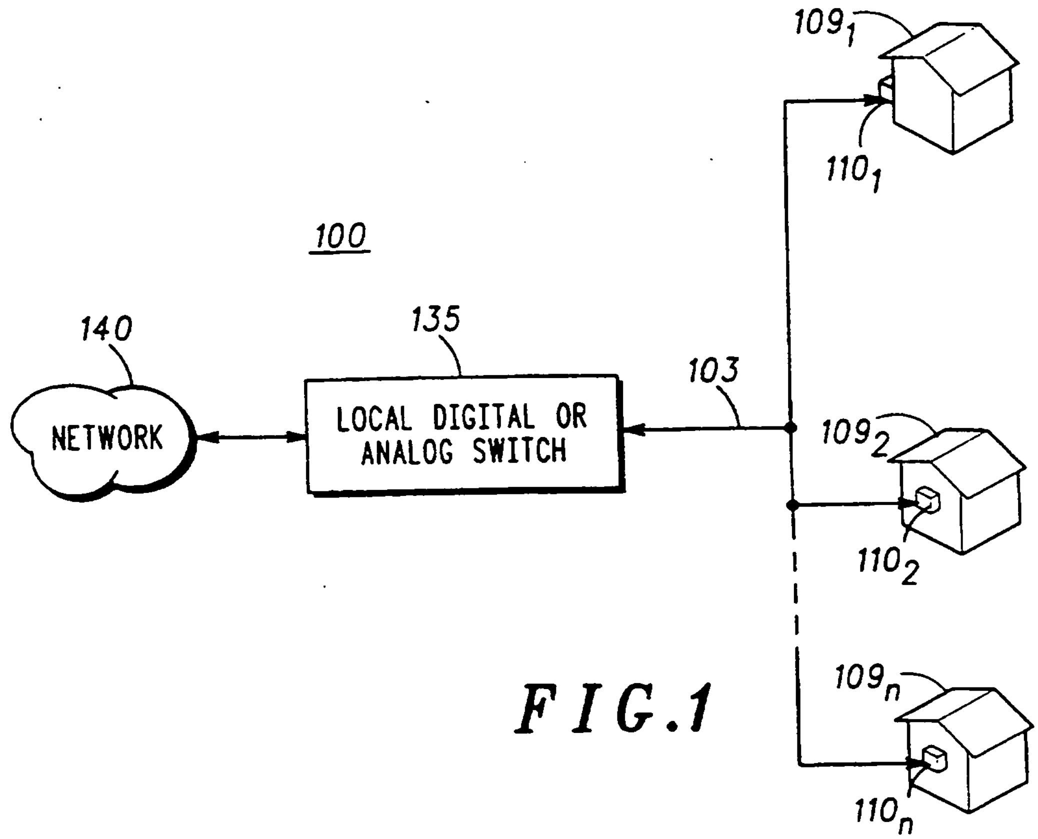

100
140
NETWORK
135
LOCAL DIGITAL OR ANALOG SWITCH
103
109 1
110 1
109 2
110 2
109 n
110 n

FIG.1

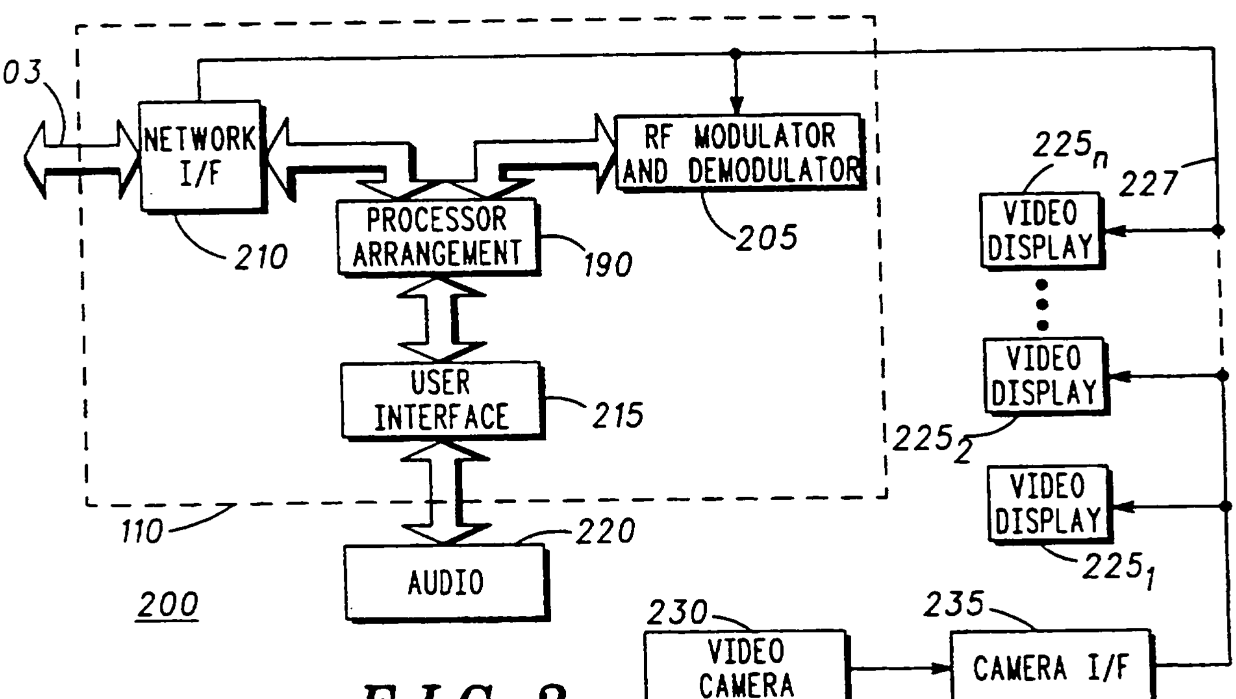

103
NETWORK I/F
210
RF MODULATOR AND DEMODULATOR
205
PROCESSOR ARRANGEMENT
190
USER INTERFACE
215
110
AUDIO
220
200
225 n
227
VIDEO DISPLAY
VIDEO DISPLAY
225 2
VIDEO DISPLAY
225 1
230
VIDEO CAMERA
235
CAMERA I/F

FIG.2

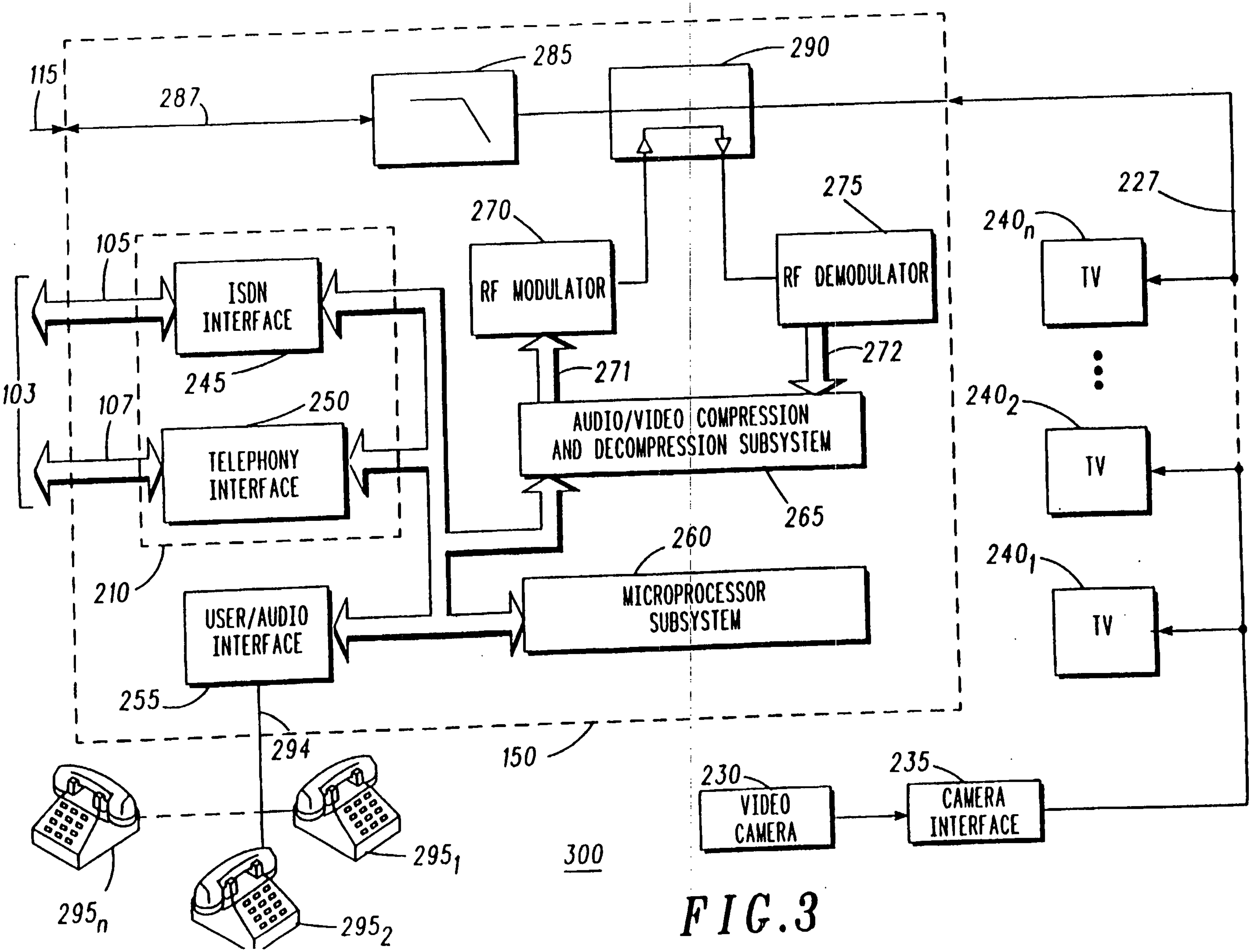

115
287
285
290
270
275
227
240n
105
ISDN INTERFACE
RF MODULATOR
RF DEMODULATOR
TV
245
271
272
103
107
250
AUDIO/VIDEO COMPRESSION AND DECOMPRESSION SUBSYSTEM
2402
TELEPHONY INTERFACE
TV
265
260
210
2401
USER/AUDIO INTERFACE
MICROPROCESSOR SUBSYSTEM
TV
255
294
150
230
235
VIDEO CAMERA
CAMERA INTERFACE
2951
300
295n
2952

FIG.3

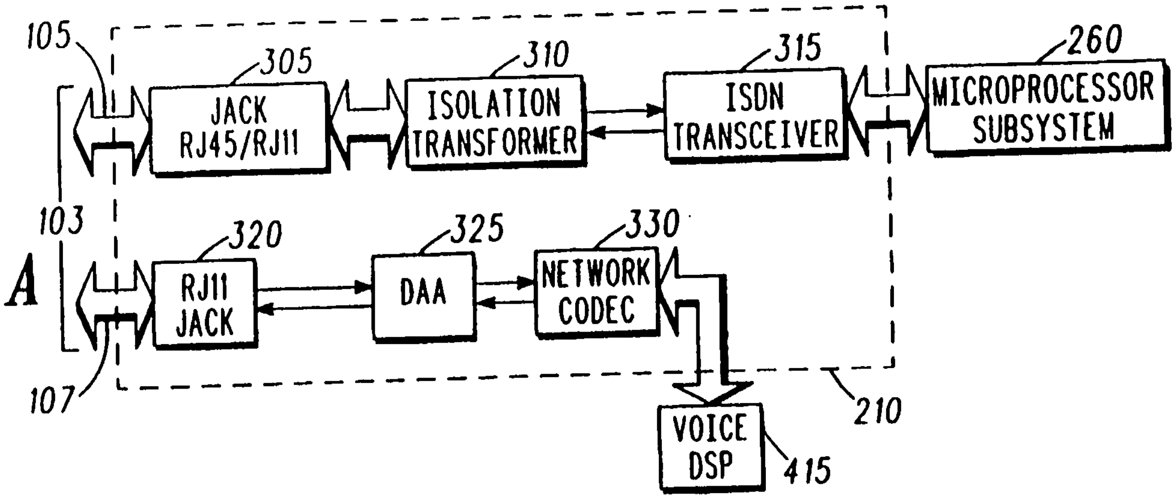


FIG.4A

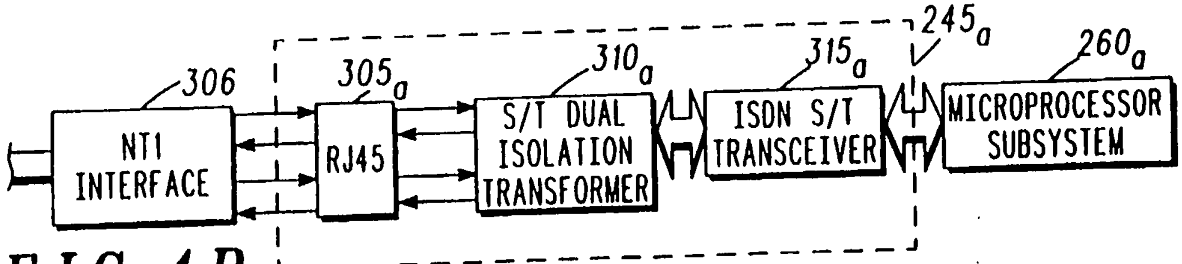


FIG.4B

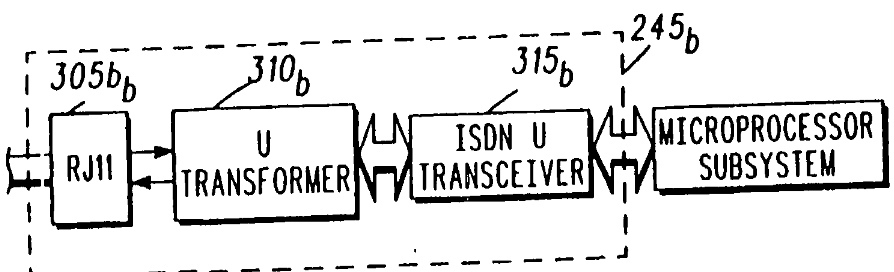
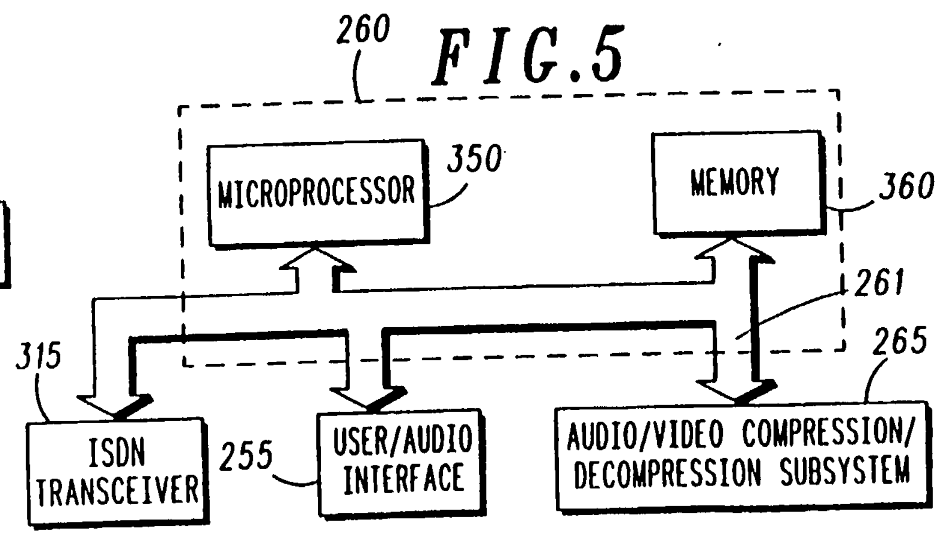


FIG.4C

FIG.5



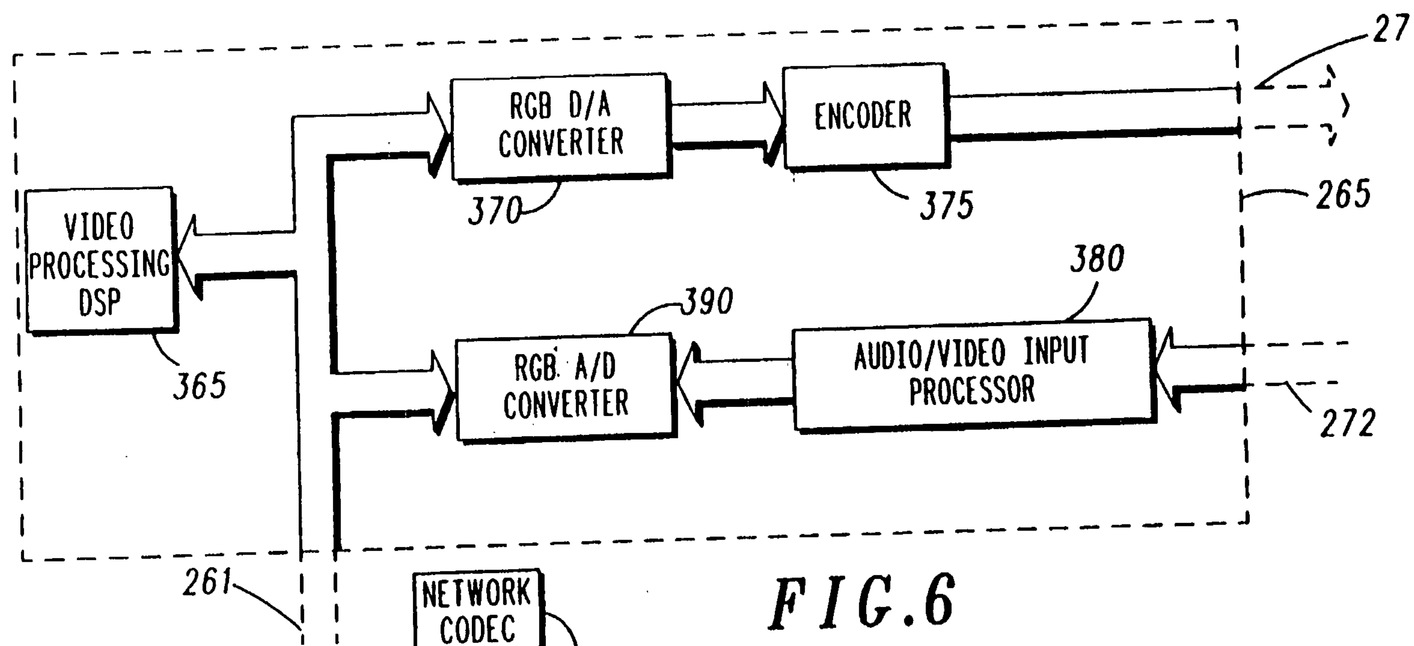


FIG.6

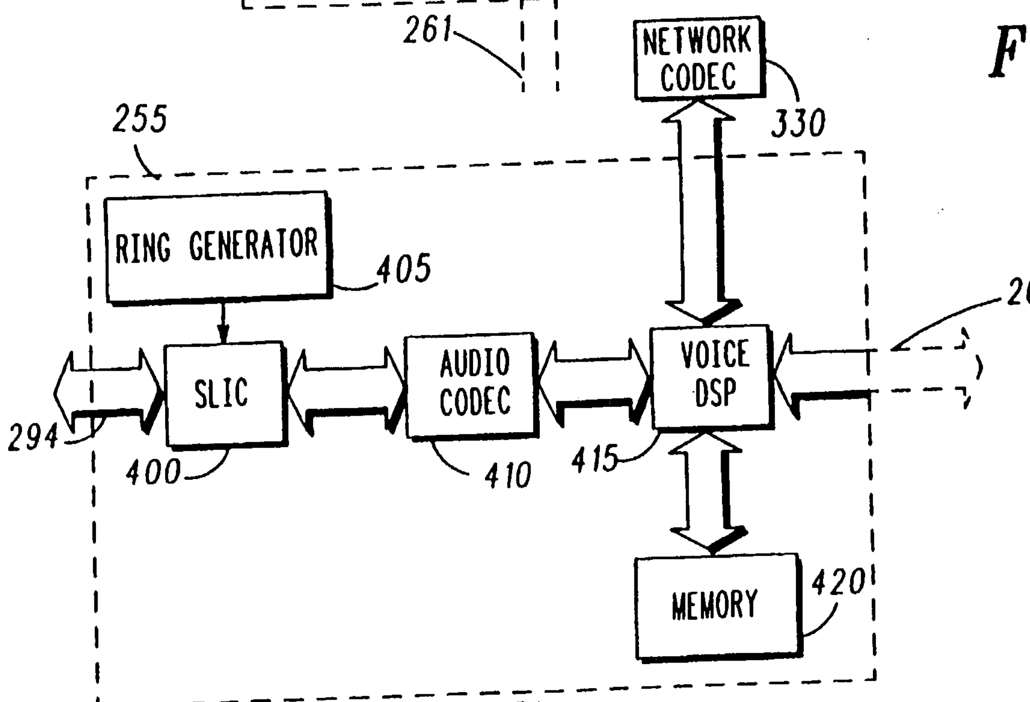


FIG.7

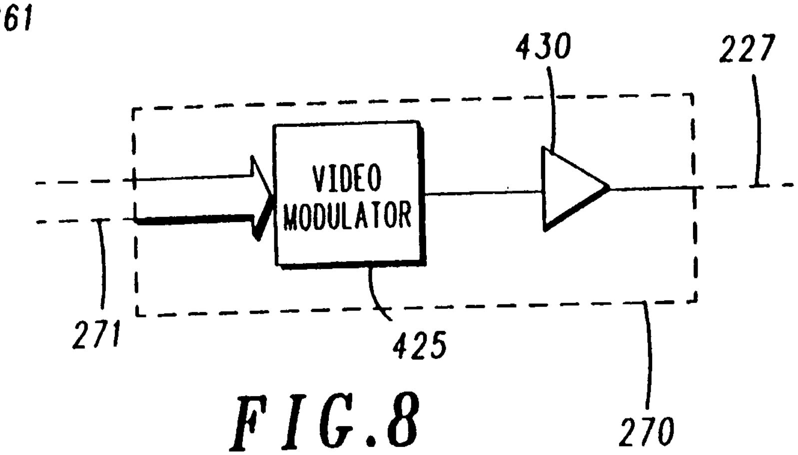
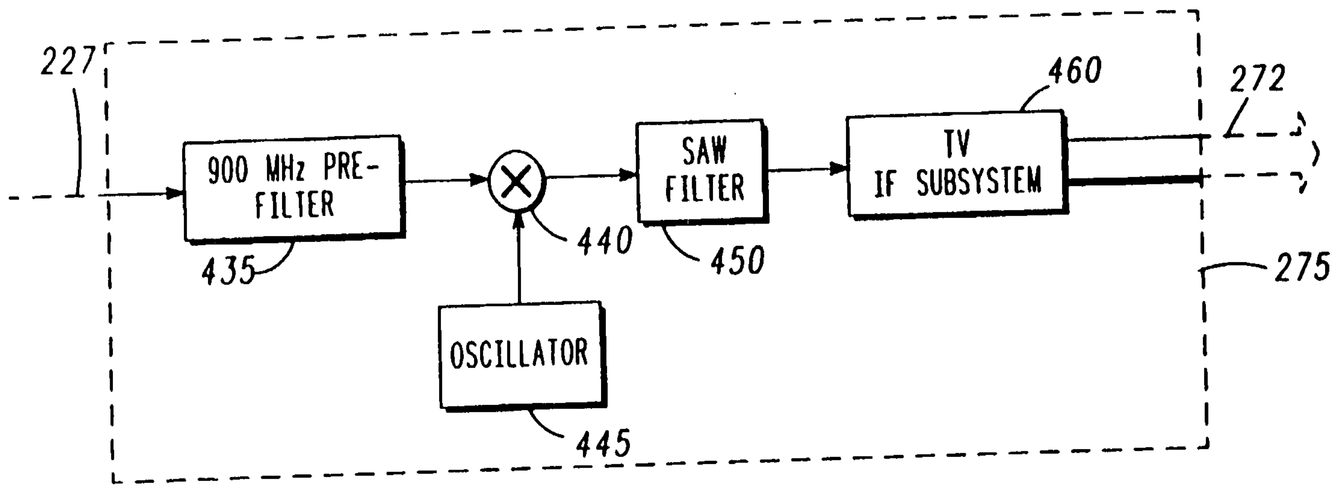
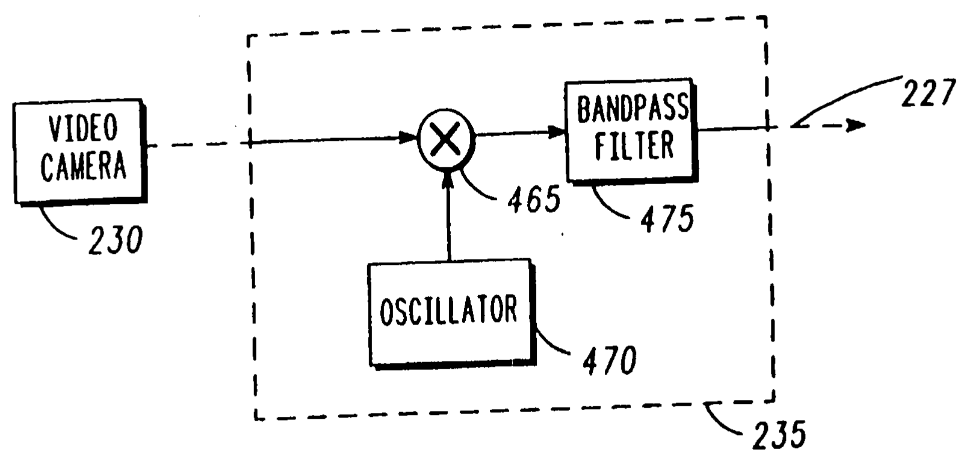


FIG.8



227
900 MHz PRE-FILTER
435
440
OSCILLATOR
445
SAW FILTER
450
TV IF SUBSYSTEM
460
272
275

FIG.9



VIDEO CAMERA
230
465
OSCILLATOR
470
BANDPASS FILTER
475
235
227

FIG.10

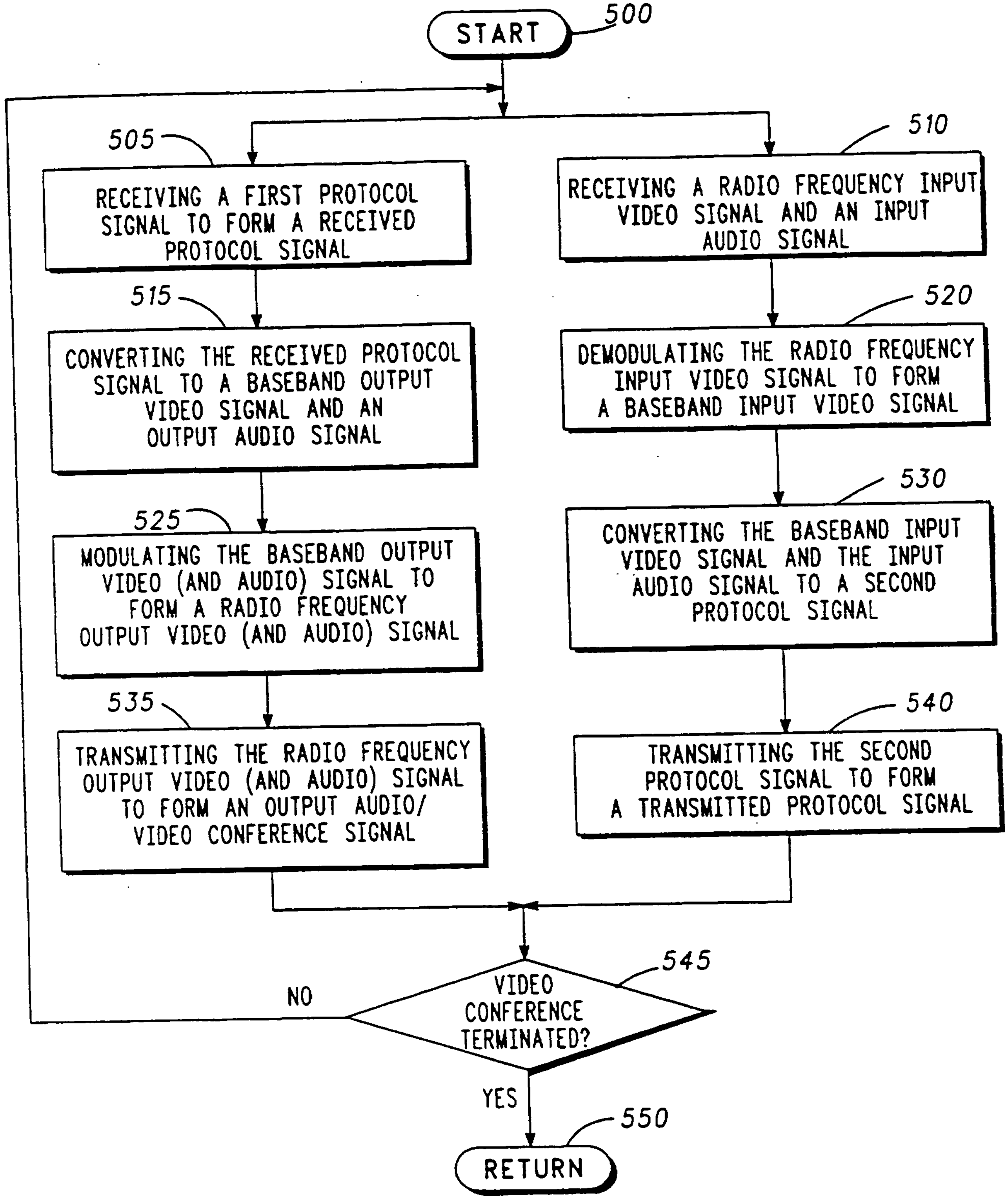

START
500
505
RECEIVING A FIRST PROTOCOL SIGNAL TO FORM A RECEIVED PROTOCOL SIGNAL
510
RECEIVING A RADIO FREQUENCY INPUT VIDEO SIGNAL AND AN INPUT AUDIO SIGNAL
515
CONVERTING THE RECEIVED PROTOCOL SIGNAL TO A BASEBAND OUTPUT VIDEO SIGNAL AND AN OUTPUT AUDIO SIGNAL
520
DEMODULATING THE RADIO FREQUENCY INPUT VIDEO SIGNAL TO FORM A BASEBAND INPUT VIDEO SIGNAL
525
MODULATING THE BASEBAND OUTPUT VIDEO (AND AUDIO) SIGNAL TO FORM A RADIO FREQUENCY OUTPUT VIDEO (AND AUDIO) SIGNAL
530
CONVERTING THE BASEBAND INPUT VIDEO SIGNAL AND THE INPUT AUDIO SIGNAL TO A SECOND PROTOCOL SIGNAL
535
TRANSMITTING THE RADIO FREQUENCY OUTPUT VIDEO (AND AUDIO) SIGNAL TO FORM AN OUTPUT AUDIO/ VIDEO CONFERENCE SIGNAL
540
TRANSMITTING THE SECOND PROTOCOL SIGNAL TO FORM A TRANSMITTED PROTOCOL SIGNAL
545
VIDEO CONFERENCE TERMINATED?
NO
YES
550
RETURN

FIG.11

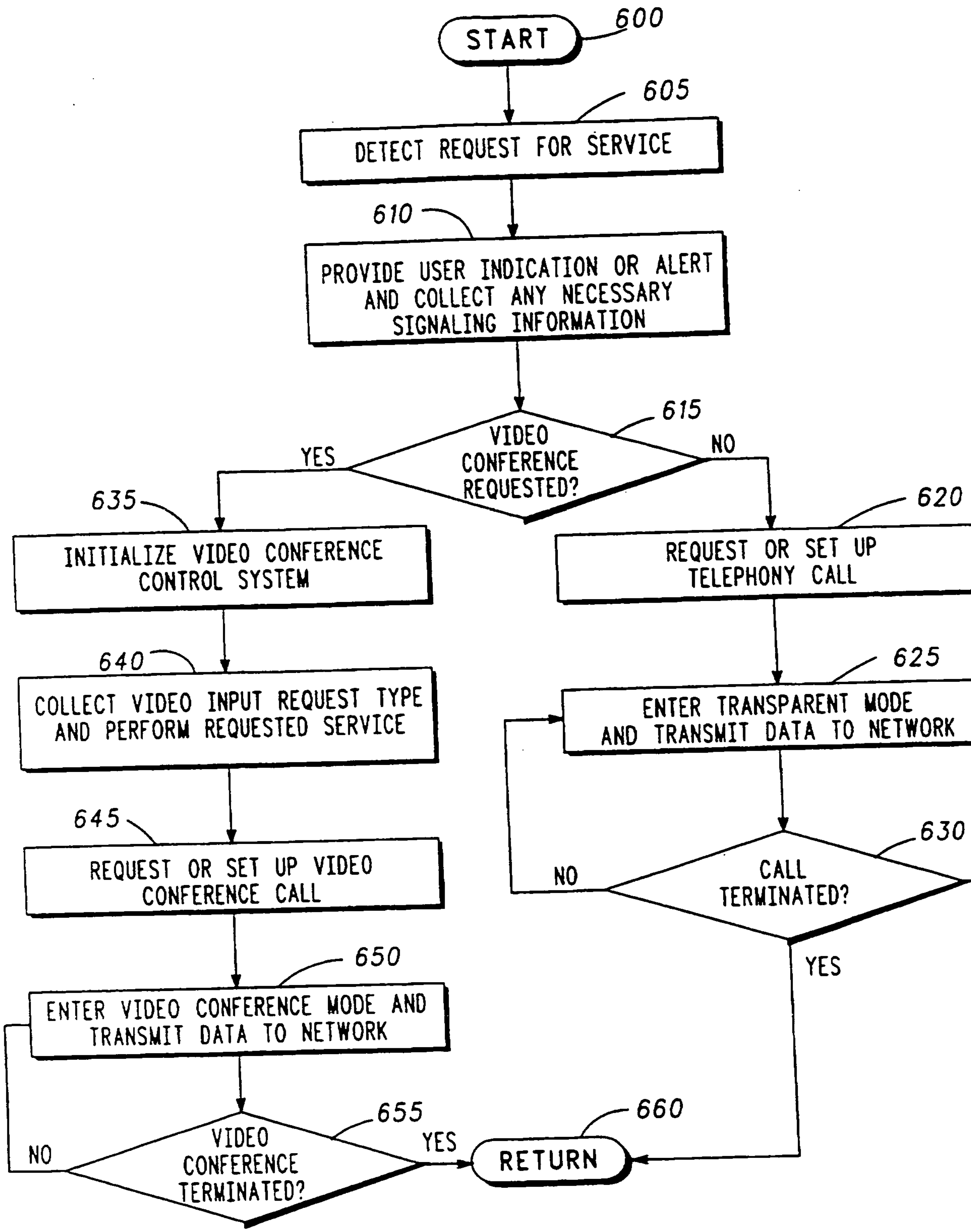

START
600
605
DETECT REQUEST FOR SERVICE
610
PROVIDE USER INDICATION OR ALERT AND COLLECT ANY NECESSARY SIGNALING INFORMATION
615
VIDEO CONFERENCE REQUESTED?
YES
NO
635
INITIALIZE VIDEO CONFERENCE CONTROL SYSTEM
620
REQUEST OR SET UP TELEPHONY CALL
640
COLLECT VIDEO INPUT REQUEST TYPE AND PERFORM REQUESTED SERVICE
625
ENTER TRANSPARENT MODE AND TRANSMIT DATA TO NETWORK
645
REQUEST OR SET UP VIDEO CONFERENCE CALL
630
NO
CALL TERMINATED?
YES
650
ENTER VIDEO CONFERENCE MODE AND TRANSMIT DATA TO NETWORK
655
NO
VIDEO CONFERENCE TERMINATED?
YES
660
RETURN

FIG.12

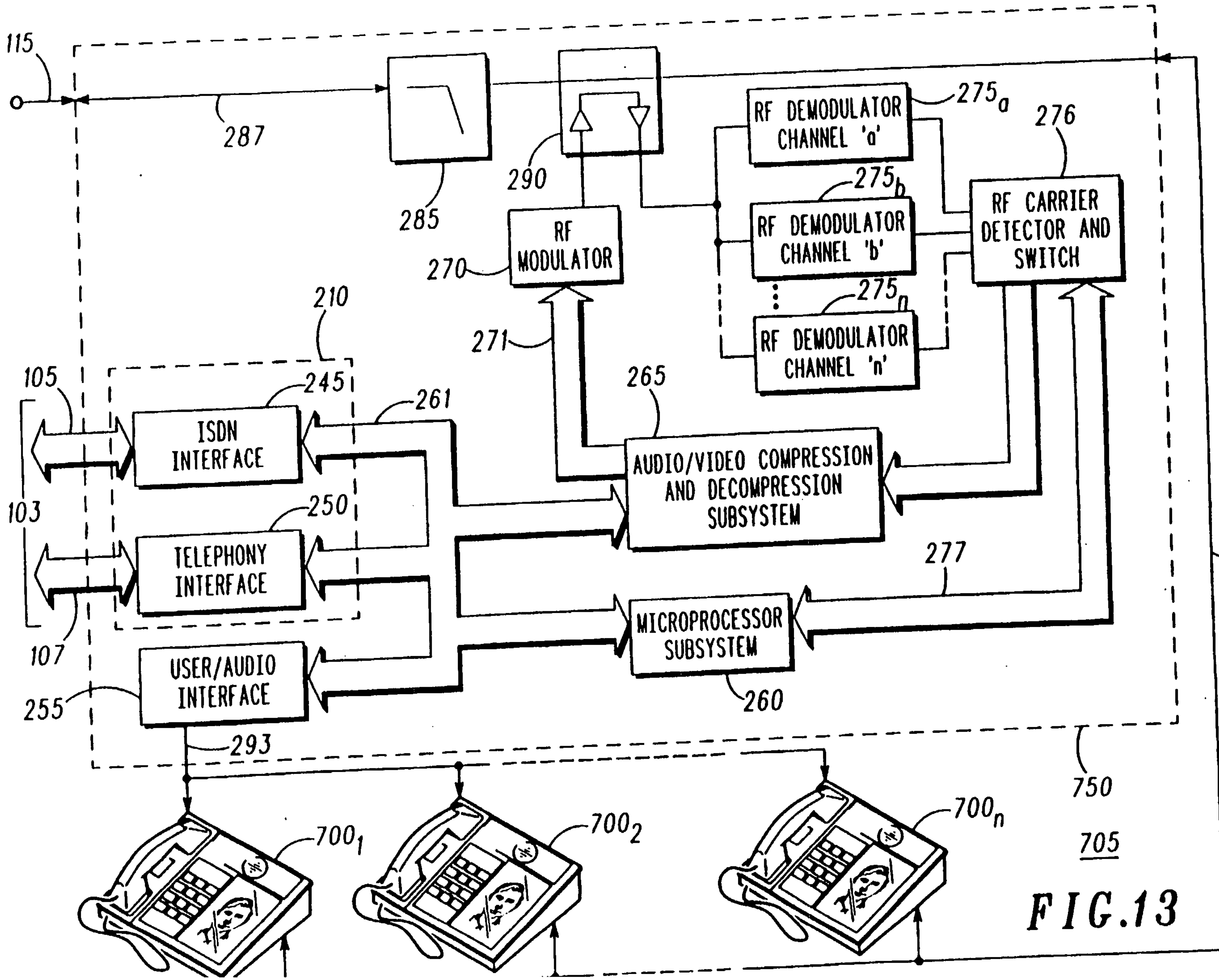

115
287
285
290
RF MODULATOR
270
271
RF DEMODULATOR CHANNEL 'a'
275_a
RF DEMODULATOR CHANNEL 'b'
275_b
RF DEMODULATOR CHANNEL 'n'
275_n
RF CARRIER DETECTOR AND SWITCH
276
210
105
103
107
ISDN INTERFACE
245
TELEPHONY INTERFACE
250
USER/AUDIO INTERFACE
255
261
265
AUDIO/VIDEO COMPRESSION AND DECOMPRESSION SUBSYSTEM
277
MICROPROCESSOR SUBSYSTEM
260
293
750
227
700_1
700_2
700_n
705

FIG.13

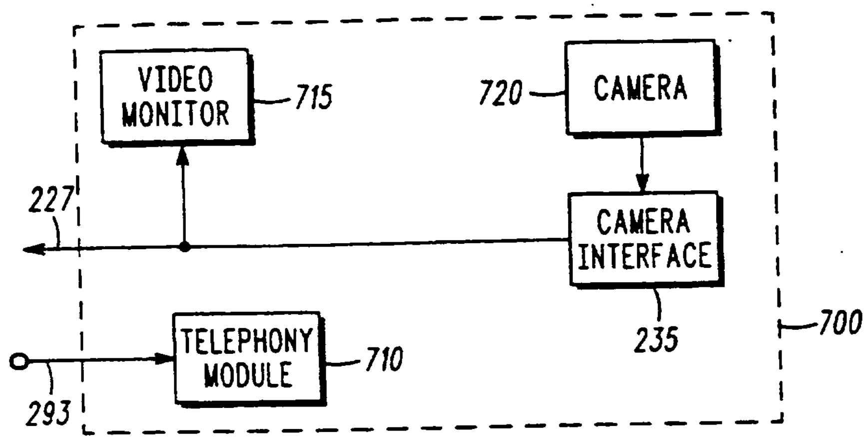

VIDEO MONITOR
715
720
CAMERA
227
CAMERA INTERFACE
TELEPHONY MODULE
710
235
700
293

FIG.14

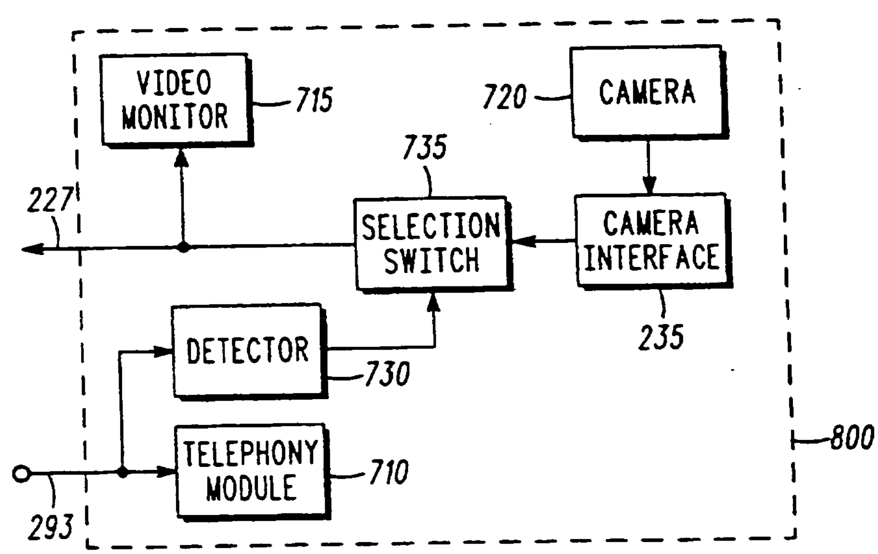

VIDEO MONITOR
715
720
CAMERA
735
227
SELECTION SWITCH
CAMERA INTERFACE
DETECTOR
730
235
TELEPHONY MODULE
710
800
293

FIG.15

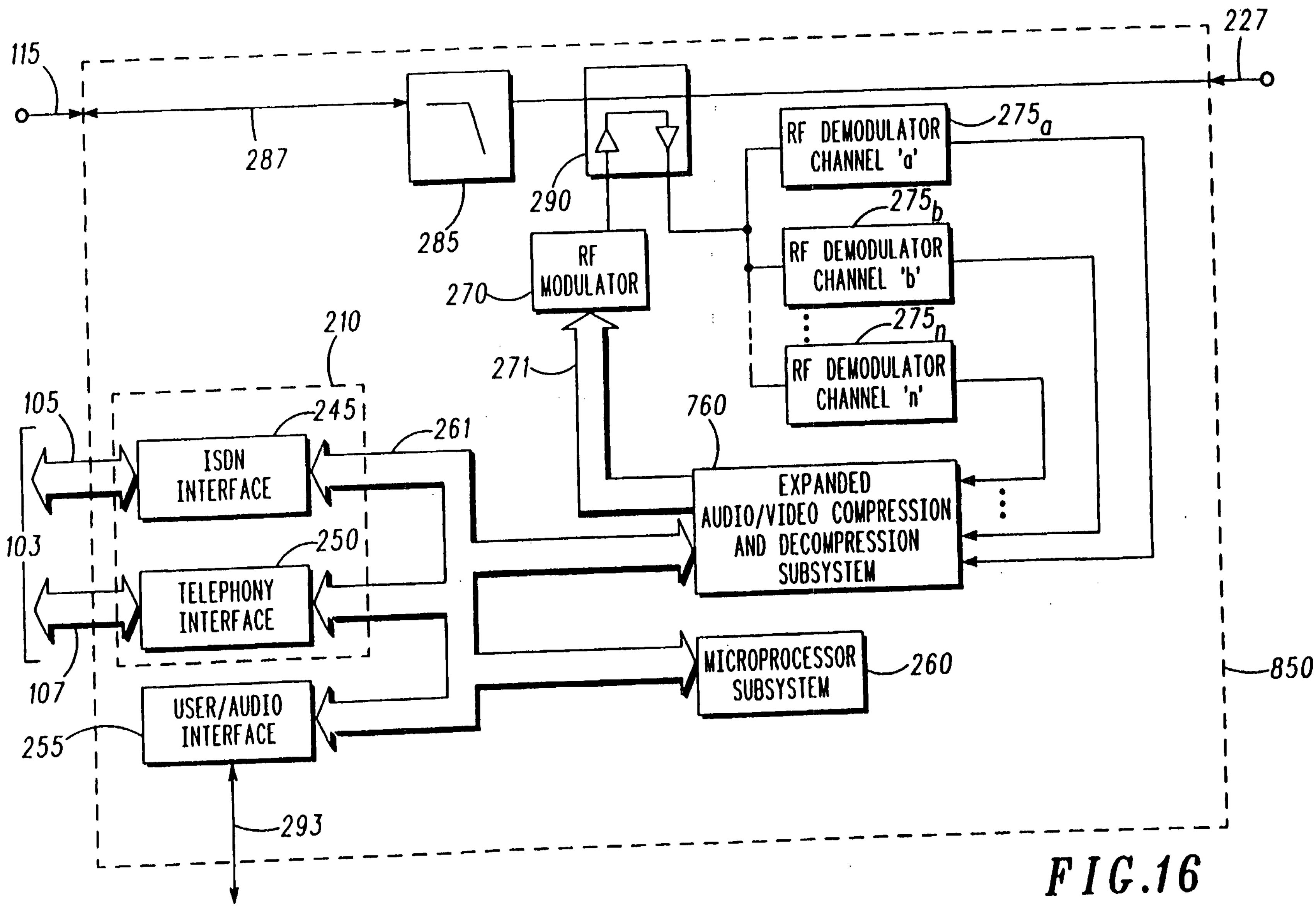
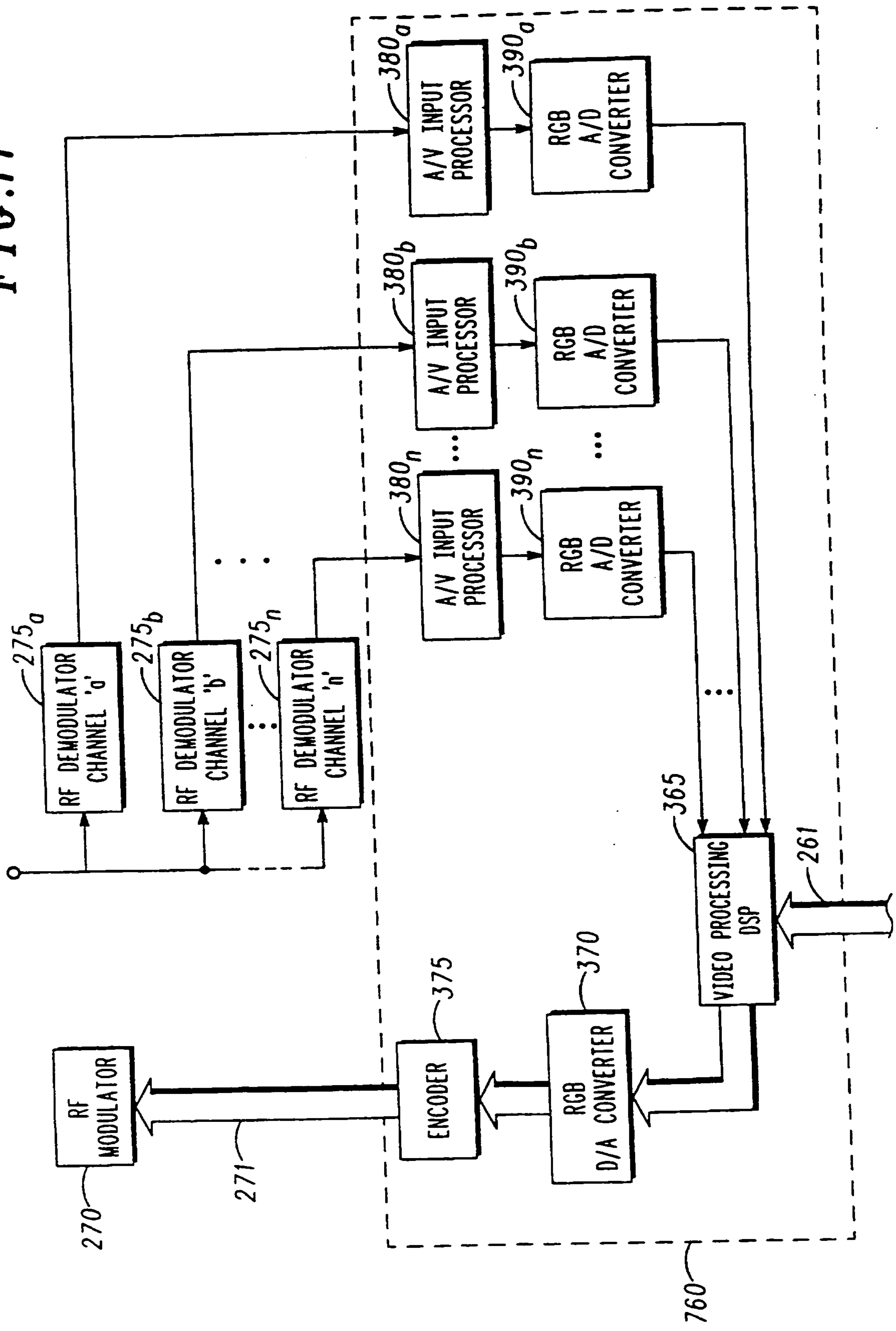

115
287
285
290
RF MODULATOR
270
271
227
RF DEMODULATOR CHANNEL 'a'
275_a
RF DEMODULATOR CHANNEL 'b'
275_b
RF DEMODULATOR CHANNEL 'n'
275_n
760
EXPANDED AUDIO/VIDEO COMPRESSION AND DECOMPRESSION SUBSYSTEM
MICROPROCESSOR SUBSYSTEM
260
850
210
245
ISDN INTERFACE
250
TELEPHONY INTERFACE
USER/AUDIO INTERFACE
255
261
105
103
107
293

FIG.16

FIG.17



RF DEMODULATOR CHANNEL 'a'
275a
RF DEMODULATOR CHANNEL 'b'
275b
RF DEMODULATOR CHANNEL 'n'
275n
A/V INPUT PROCESSOR
380a
380b
380n
RGB A/D CONVERTER
390a
390b
390n
VIDEO PROCESSING DSP
365
261
RGB D/A CONVERTER
370
ENCODER
375
RF MODULATOR
270
271
760

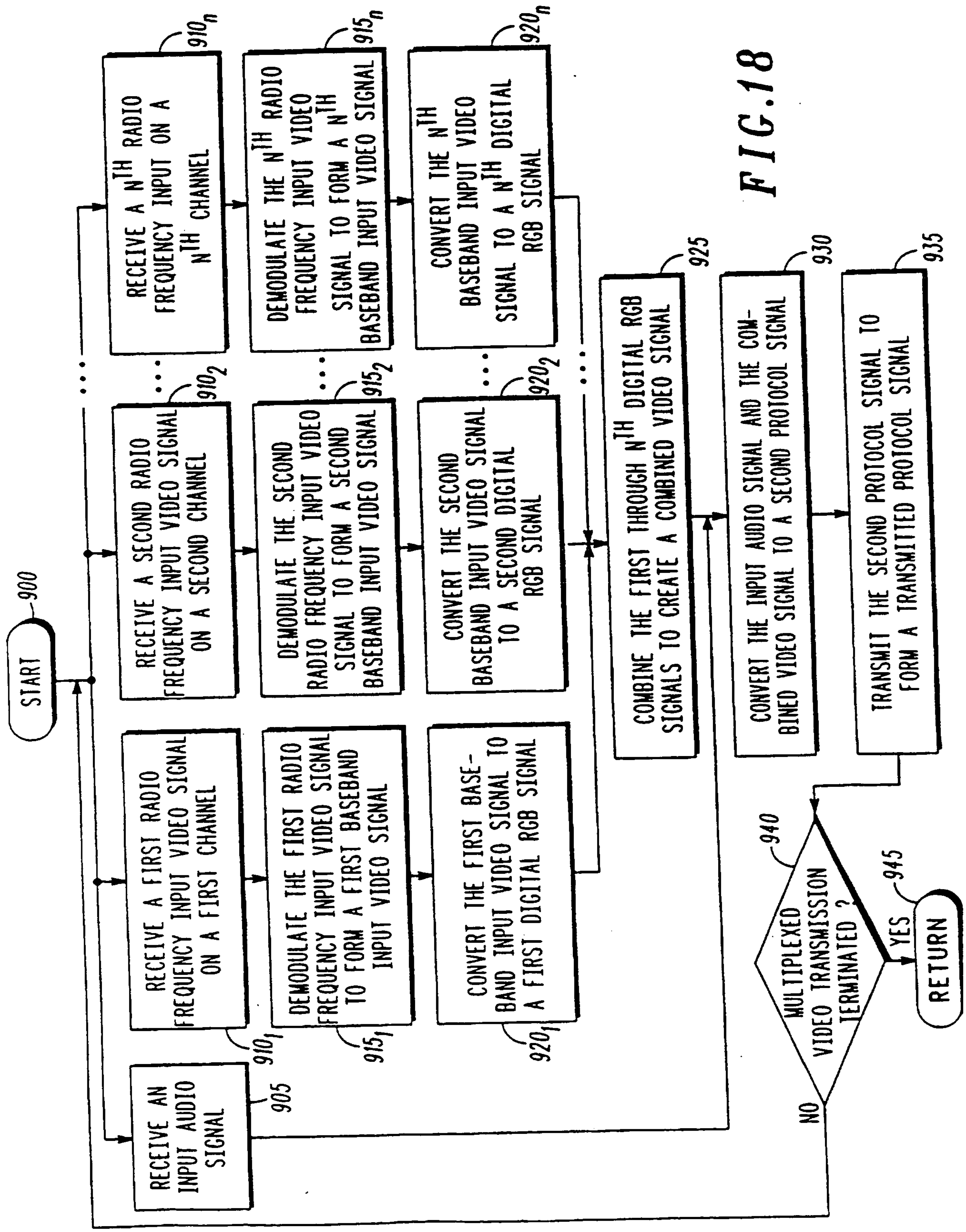

START
900
RECEIVE AN INPUT AUDIO SIGNAL
905
RECEIVE A FIRST RADIO FREQUENCY INPUT VIDEO SIGNAL ON A FIRST CHANNEL
910_1
RECEIVE A SECOND RADIO FREQUENCY INPUT VIDEO SIGNAL ON A SECOND CHANNEL
910_2
RECEIVE A N^{TH} RADIO FREQUENCY INPUT ON A N^{TH} CHANNEL
910_n
DEMODULATE THE FIRST RADIO FREQUENCY INPUT VIDEO SIGNAL TO FORM A FIRST BASEBAND INPUT VIDEO SIGNAL
915_1
DEMODULATE THE SECOND RADIO FREQUENCY INPUT VIDEO SIGNAL TO FORM A SECOND BASEBAND INPUT VIDEO SIGNAL
915_2
DEMODULATE THE N^{TH} RADIO FREQUENCY INPUT VIDEO SIGNAL TO FORM A N^{TH} BASEBAND INPUT VIDEO SIGNAL
915_n
CONVERT THE FIRST BASEBAND INPUT VIDEO SIGNAL TO A FIRST DIGITAL RGB SIGNAL
920_1
CONVERT THE SECOND BASEBAND INPUT VIDEO SIGNAL TO A SECOND DIGITAL RGB SIGNAL
920_2
CONVERT THE N^{TH} BASEBAND INPUT VIDEO SIGNAL TO A N^{TH} DIGITAL RGB SIGNAL
920_n
COMBINE THE FIRST THROUGH N^{TH} DIGITAL RGB SIGNALS TO CREATE A COMBINED VIDEO SIGNAL
925
CONVERT THE INPUT AUDIO SIGNAL AND THE COMBINED VIDEO SIGNAL TO A SECOND PROTOCOL SIGNAL
930
TRANSMIT THE SECOND PROTOCOL SIGNAL TO FORM A TRANSMITTED PROTOCOL SIGNAL
935
MULTIPLEXED VIDEO TRANSMISSION TERMINATED ?
940
YES
NO
RETURN
945

FIG.18

2318021

VIDEOPHONE APPARATUS, METHOD AND SYSTEM FOR WIRELINE AUDIO AND VIDEO CONFERENCING AND TELEPHONY

Cross-Reference to Related Applications

This application is related to the following United States Patent Applications, each incorporated by reference herein, with priority claimed for all commonly disclosed subject matter:

Newlin et al., United States Patent Application Serial No. 08/658,792, filed June 5, 1996, entitled "Audio/Visual Communication System and Method Thereof", Motorola Docket No. PD05634AM (the "first related application");

Burke et al., United States Patent Application Serial No. 08/706,100, filed August 30, 1996, entitled "Apparatus, Method And System For Audio And Video Conferencing And Telephony", Motorola Docket No. PD05686AM (the "second related application");

Burke et al., United States Patent Application Serial No. 08/715,887, filed September 18, 1996, entitled "Videophone Apparatus, Method And System For Audio And Video Conferencing And Telephony", Motorola Docket No. PD05689AM (the "third related application");

Newlin et al., United States Patent Application Serial No. 08/672,819, filed June 28, 1996, entitled "ISDN Communication System and Method Thereof", Motorola Docket No. PD05677AM (the "fourth related application"); and

Burke et al., United States Patent Application Serial No. 08/725,602, filed October 3, 1996, entitled "Apparatus, Method And System For Wireline Audio And Video

Conferencing And Telephony", Motorola Docket No. PD05703AM (the "fifth related application").

Field of the Invention

This invention relates in general to audio and video communications systems and, more specifically, to an videophone apparatus, method and system for wireline audio and video conferencing and telephony.

Background of the Invention

Currently, audio and video (visual) conferencing capabilities are implemented as computer based systems, such as in personal computers ("PCs"), as stand-alone, "roll about" room systems, and as videophones. These systems typically require new and significant hardware, software and programming, and may also require significant communications network connections, for example, multiple channels ("DSOs") of an Integrated Services Digital Network ("ISDN") connection or a T1/E1 connection.

For example, stand-alone, "roll about" room systems for audio and video conferencing typically require dedicated hardware at significant expense, in the tens of thousands of dollars, utilizing dedicated video cameras, television or video displays, microphone systems, and the additional video conferencing equipment. Such systems may also require as many as six (or more) contiguous ISDN B channels (or T1/E1 DSOs), each operating at 64 kbps (kilobits per second). Such communication network capability is also expensive and potentially unnecessary, particularly when the additional channels are not in continuous use.

Current audio/visual telephony or conferencing systems are also limited to providing such audio/visual

functionality only at designated nodes, i.e., the specific system location, and are neither mobile nor distributed (having multiple locations). Stand-alone, "roll about" room systems allow such audio and video conferencing only within or at that particular physical location. Videophones are also currently limited to their installed locations. Similarly, PC based systems provide such functionality only at the given PC having the necessary network connections (such as ISDN) and having the specified audio/visual conferencing equipment, such as a video camera, microphone, and the additional computer processing boards which provide for the audio/visual processing . For other PCs to become capable of such audio/visual conferencing functionality, they must also be equipped with any necessary hardware, software, programming and network connections.

Such conventional audio/visual conferencing systems are also difficult to assemble, install, and use. For example, the addition of audio/visual functionality to a PC requires the addition of a new PC card, camera, microphone, the installation of audio/visual control software, and the installation of new network connections, such as ISDN. PC based systems typically require, at a minimum, ISDN basic rate interface service, consisting of 2 ISDN B channels (each operating at 64 kbps) plus one D channel (operating at 16 kbps). In addition, such network connectivity may require additional programming of the PC with necessary ISDN specific configuration information, such as configuration information specific to the central office switch type of the service provider and ISDN service profile identifier (SPID) information. Video conference call set up procedures typically are also difficult and complicated utilizing these current systems.

Conventional audio/visual telephony and conferencing equipment is also limited to communication

with similar equipment at the far end (remote location). For example, videophone systems which utilize typical telephone systems ("POTS" (plain old telephone service)) transmit information in analog form, for example, as trellis code modulated data, at V.34 and V.34bis rates (e.g., highest data rates of approximately 28.8 to 33 kbps). Such POTS-based videophone systems would not be compatible with ISDN audio/visual conferencing and telephony systems which transmit information in digital form, such as utilizing Q.931 message signaling, Q.921 LAPD datalink, and Q.910 physical interface digital protocols, with data rates of 128 kbps (two B channels) or more (with additional channels or DSOs).

In addition, such current audio/visual telephony and conferencing equipment are relatively expensive and, in most instances, sufficiently expensive to be prohibitive for in-home or other consumer use. For example, the cost of roll about, room based systems is typically tens of thousands of dollars. PC based videoconferencing systems are also expensive, with costs in the thousands of dollars.

Current audio/visual telephony and conferencing equipment also do not provide for multiple, simultaneous video conferences from more than one location. In addition, current systems (such as those in PCs) do not provide for multiplexed video conference sessions, in which the output video may include display of video input from several video cameras at multiple locations.

Accordingly, a need has remained for audio/visual conferencing and telephony systems, equipment, and methods which may operate at more than one designated node or location within the user premises, or may be mobile, or may be configured as needed for additional locations. Such a system should be compatible for use with other existing video conferencing systems, should be user friendly, easy to install and use, and should be relatively less expensive for in-

home purchase and use by consumers. In addition, such a system should be able to provide multiple video conferencing sessions which may originate from multiple locations.

Brief Description of the Drawings

FIG. 1 is a block diagram illustrating an audio/video network configuration for a video access apparatus in accordance with the invention disclosed in the fifth related application.

FIG. 2 is a high level block diagram illustrating a first embodiment of a video access apparatus and a first embodiment of a video conferencing system in accordance with the invention disclosed in the fifth related application.

FIG. 3 is a detailed block diagram illustrating a second embodiment of a video access apparatus and a second embodiment of a video conferencing system in accordance with the invention disclosed in the fifth related application.

FIG. 4A is a block diagram illustrating a wireline network interface of the preferred apparatus embodiment in accordance with the invention disclosed in the fifth related application.

FIG. 4B is a block diagram illustrating an embodiment of an ISDN portion of a wireline network interface utilizing an ISDN S/T interface in accordance with the invention disclosed in the fifth related application.

FIG. 4C is a block diagram illustrating an embodiment of an ISDN portion of a wireline network interface utilizing an ISDN U interface in accordance with the invention disclosed in the fifth related application.

FIG. 5 is a block diagram illustrating a microprocessor subsystem of the preferred apparatus embodiment in accordance with the invention disclosed in the fifth related application.

FIG. 6 is a block diagram illustrating a audio/video compression and decompression subsystem of the preferred apparatus embodiment in accordance with the invention disclosed in the fifth related application.

FIG. 7 is a block diagram illustrating a user audio interface of the preferred apparatus embodiment in accordance with the invention disclosed in the fifth related application.

FIG. 8 is a block diagram illustrating an RF modulator of the preferred apparatus embodiment in accordance with the invention disclosed in the fifth related application.

FIG. 9 is a block diagram illustrating an RF demodulator of the preferred apparatus embodiment in accordance with the invention disclosed in the fifth related application.

FIG. 10 is a block diagram illustrating a camera interface of the preferred apparatus embodiment in accordance with the invention disclosed in the fifth related application.

FIG. 11 is a flow diagram illustrating the method of the preferred embodiment in accordance with the invention disclosed in the fifth related application.

FIG. 12 is a flow diagram illustrating the telephony and video conference control methodology of the preferred embodiment in accordance with the invention disclosed in the fifth related application.

FIG. 13 is a block diagram illustrating a third embodiment of a video access apparatus 750 and a third embodiment of a video conferencing system 705, utilizing a plurality of videophone apparatuses 700, in accordance with the present invention.

FIG. 14 is a detailed block diagram illustrating a first embodiment of a videophone apparatus 700 in accordance with the present invention.

FIG. 15 is a detailed block diagram illustrating a second embodiment of a videophone apparatus 800 in accordance with the present invention.

FIG. 16 is a detailed block diagram illustrating a fourth embodiment of a video access apparatus 850 in accordance with the present invention.

FIG. 17 is a block diagram illustrating an expanded audio/video compression and decompression subsystem 760 in accordance with the present invention.

FIG. 18 is a flow diagram illustrating the video multiplexing method in accordance with the present invention.

Detailed Description of the Invention

As mentioned above, a need has remained for audio/visual conferencing and telephony systems, apparatus, and methods which may operate at more than one designated node or location within user premises, or may be mobile, or may be configured as needed for additional locations. As illustrated in FIGs. 1 through 12 discussed below, the preferred embodiment of the invention provides for such audio and visual conferencing and telephony capability at one or more locations within the user premises, may be mobile, and may be configured as needed for additional locations. In addition, in accordance with the preferred embodiment, the audio/visual conferencing and telephony system utilizes equipment typically found in consumers' homes or premises, such as existing televisions, video cameras or camcorders, and telephones. In addition, such a system is designed to be compatible for use with other existing video conferencing systems, may be utilized over a variety of connected telecommunications networks (such as ISDN or POTS), is

user friendly, easy to install and use, and should be relatively less expensive for in-home purchase and use by consumers.

Also as mentioned above, a need has also remained for audio/visual telephony and conferencing equipment to provide for multiple, simultaneous video conferences from more than one location. In accordance with the preferred embodiment of the present invention, multiple, simultaneous video conferences from more than one location may occur, providing for multiplexed video conference sessions, in which the output video may include display of video input from several video cameras at multiple locations. In addition, in accordance with the present invention, various videophone apparatus embodiments are disclosed, which provide an alternative to the use of telephones, video cameras and televisions as utilized and disclosed in the invention of the fifth related application. The various videophone apparatus embodiments in accordance with the present invention also may be utilized over a variety of connected telecommunications networks (such as ISDN or POTS), are user friendly, easy to install and use, and also should be relatively less expensive for in-home purchase and use by consumers.

FIG. 1 is a block diagram illustrating a configuration of an audio/video network 100 for a video access apparatus 110 in accordance with the invention disclosed in the fifth related application. As illustrated in FIG. 1, video access apparatus 110_1 through video access apparatus 110_n (individually and collectively referred to as a video access apparatus(es) 110) may have an outdoor location, for example, at subscriber premises 109_1 (video access apparatus 110_1), or may have indoor locations, for example, at subscriber premises 109_2 and 109_n (video access apparatus 110_1 and video access apparatus 110_n). The video access apparatus 110 illustrated in FIG. 1 may have a first embodiment as illustrated in FIG. 2, or a

second and preferred embodiment as video access apparatus 150 illustrated in FIG. 3, and as a consequence, as used herein, reference to any of the embodiments of the video access apparatuses 110 or 150 shall be understood to mean and include the other apparatus embodiment or its equivalents. Referring to FIG. 1, in accordance with the invention, the video access apparatus 110 provides audio and video telephony and conferencing services over a first communication channel 103 which, in the preferred embodiment, is wireline, such as one or more twisted pair wires, utilized in the audio/video network 100 (which may have multiple configurations). Also in the preferred embodiment, the first communication channel 103 may be utilized for both digital and analog communications, such as ISDN and ordinary telephony commonly known as POTS. The first communication channel 103, in turn, is connected through a local digital (or analog) switch 135 to a network 140. The network 140, for example, may be a public switched telephone network ("PSTN") or an Integrated Services Digital Network ("ISDN"), or any combination of such existing or future telecommunications networks.

As illustrated above, the video access apparatus 110 (or 150) of the invention disclosed in the fifth related application is directly coupleable (through a local digital or analog switch 135 of a network provider central office) to a network 140, such as ISDN or PSTN. As a consequence, the video access apparatus 110 (or 150) of the invention disclosed in the fifth related application may be utilized with currently existing telecommunications infrastructure, such as ISDN or PSTN. In contrast, the video access apparatus disclosed in the second and third related applications communicated with an intervening primary station which then provided access both to a cable video services infrastructure and to a network, such as ISDN or PSTN, utilizing a protocol such as CACS (Cable

ACcess Signaling) over a communication channel (such as a preferred hybrid fiber coaxial cable). While use of CACS and the system disclosed in the second and third related applications have certain advantages, such as very high speed, low error rate, asynchronous packet data transfer with very high data throughput, utilizing on demand channel assignment, direct network connectivity was precluded. As a consequence, in accordance with the present invention and the invention disclosed in the fifth related application, direct network connectivity is provided, such that the video access apparatus 110 (or 150) may be utilized, for video conferencing and telephony, directly with currently existing telecommunications network infrastructure, such as ISDN or PSTN, without further infrastructure requirements.

FIG. 2 is a high level block diagram illustrating a first embodiment of a video access apparatus, namely, video access apparatus 110, and illustrating a video conferencing system 200, in accordance with the invention disclosed in the fifth related application. The video conferencing system 200, in accordance with the invention disclosed in the fifth related application, includes a video access apparatus 110, audio 220, one or more video displays 225_1 through 225_n (individually and collectively referred to as video display(s) 225), camera interface 235, and video camera 230. The video access apparatus 110 is coupleable to a first communication channel 103, for communication with a network 140 via a local digital or analog switch 135, and is coupled to a second communication channel 227, typically located within or about the user (or subscriber) premises 109. For example, the second communication channel 227 may be an internal 75 Ohm coaxial cable typically utilized with cable television, or may be another form of communication channel, such as twisted pair or other wireline, wireless, or PLC (power line carrier, over existing premise AC power lines). The audio 220 is coupled to

the video access apparatus 110, and may include a microphone and speaker or, as discussed below with reference to FIG. 3, may be preferably embodied as a telephone (or, equivalently, a speaker phone). One or more video displays 225 are utilized to display the incoming video portion of an audio and video conferencing call or session (incoming in the sense of having been transmitted to the video access apparatus 110 from another location), may also include a speaker for output of the incoming audio portion of an audio and video conferencing call or session, and are implemented utilizing one or more televisions in the preferred embodiment. The video camera 230 is utilized to generate the outgoing video portion of an audio and video conferencing call or session (outgoing in the sense of being transmitted from the video access apparatus 110 to another location), may also include a microphone for generation of the outgoing audio portion of an audio and video conferencing call or session, and is implemented utilizing an ordinary video camera or camcorder in the preferred embodiment. The camera interface 235 is utilized to modulate the video output signal from the video camera 230 for transmission on the second communication channel 227 to the video access apparatus 110 and, as discussed in greater detail below, the camera interface 235 also may be directly incorporated within the video camera 230.

Continuing to refer to FIG. 2, the video access apparatus 110 includes a wireline network interface 210 (also referred to equivalently as a network interface 210), a radio frequency (RF) modulator and demodulator 205 (also referred to as an RF modulator/demodulator 205), a user interface 215, and a processor arrangement 190. The wireline network interface 210 is coupleable to the first communication channel 103 for reception of a first protocol signal from the network 140, to form a received protocol signal, and for transmission of a

second protocol signal to the network 140, to form a transmitted protocol signal. These first and second protocol signals may have multiple layers and types of protocol encoding and modulation. First, such first and second protocol signals preferably include audio/video compression (and decompression) encoding (and decoding), preferably utilizing the International Telecommunications Union (ITU) H.32x series or family of protocols, such as H.320 utilized with digital services (ISDN), H.324 utilized with analog services (PSTN), H.323 utilized with LANs (local area networks), other H.32x protocols (such as H.321 and H.322), and other ITU protocols pertaining to audio/video and other data communication. In addition, in the preferred embodiment, additional protocol layers are employed, involving further encoding/decoding and/or modulation/demodulation of an H.32x encoded audio/video signal. In the preferred embodiment, for ISDN transmission and reception, ISDN protocols are utilized for encoding, decoding, framing, etc., of an H.32x encoded audio/video signal, utilizing, for example, Q.931 message signaling, Q.921 LAPD data link, and Q.910 physical layer (interface) digital protocols. Also in the preferred embodiment, for PSTN (POTS) transmission and reception, an H.32x encoded audio/video signal is further protocol encoded/decoded and modulated/demodulated utilizing the ITU V.x family or series of analog transmission protocols, such as V.34, V.34bis, or potentially or proposed higher data rate analog protocols. For example, for an analog POTS transmission, the audio/video data may be compressed and formatted utilizing ITU H.323 or H.324 protocols, then further encoded and modulated utilizing ITU V.34 or V.34bis protocols. As discussed in greater detail below with reference to FIG. 4, the wireline network interface 210 is utilized to transmit and receive analog or digital video and audio information and data (generally referred to as data), in any

given format, protocol, or modulation scheme compatible with the network 140 and any particular network connections. For example, when coupled to an ISDN via the first communication channel 103, the wireline network interface 210 will transmit and receive data in accordance with the ISDN series of protocols, such as the Q.x series.

Also, as used herein, input and output directions are defined to avoid confusion between incoming and outgoing signals because, for example, an incoming signal to the video access apparatus 110 from the network 140 will also be an outgoing signal from the video access apparatus 110 when transmitted to a video display 225 on the second communication channel 227. As a consequence, as used herein, input and output directions are defined at the interface between the video access apparatus 110, on the one hand, and the second communication channel 227 or audio 220, on the other hand, as follows: an input signal, such as an input video or audio signal, is input to the video access apparatus 110 from the second communication channel 227 (or, in the case of input audio, from the audio 220), and may originate, for example, from the video camera 230, and will be transmitted from the video access apparatus 110 to the network 140; conversely, an output signal, such as an output video or audio signal, is output from the video access apparatus 110 to the second communication channel 227 (or, in the case of output audio, to the audio 220), and may originate, for example, from a remote location via the network 140, is received by the video access apparatus 110 via the first communication channel 103, and will be transmitted or output by the video access apparatus 110 on the second communication channel 227 to a video display 225 or output to audio 220.

Continuing to refer to FIG. 2, the RF modulator and demodulator 205 is utilized to convert a baseband output

video signal (from the processor arrangement 190) to a radio frequency output video signal, for transmission on the second communication channel 227 and reception by one or more of the video displays 225, and to convert a radio frequency input video signal (from the camera interface 235) to a baseband input video signal, for input to the processor arrangement 190. The user interface 215 is utilized for reception of a control signal of a plurality of control signals, such as a request to place a telephony call, a request to place an audio and video conference call, and other control signals such as alerting signals of incoming telephony or audio and video conference calls. The processor arrangement 190 is coupled to the wireline network interface 210, to the radio frequency modulator/demodulator 205 and to the user interface 215. As explained in greater detail below, the processor arrangement 190 may be comprised of a single integrated circuit ("IC"), or may include a plurality of integrated circuits or other components connected or grouped together, such as microprocessors, digital signal processors, ASICs, associated memory (such as RAM and ROM), and other ICs and components. As a consequence, as used herein, the term processor arrangement should be understood to equivalently mean and include a single processor, or arrangement of processors, microprocessors, controllers, or some other grouping of integrated circuits which perform the functions discussed in greater detail below. For example, in the preferred embodiment, the processor arrangement 190 is implemented as illustrated in FIG. 3, and includes an audio/video compression and decompression subsystem 265 and a microprocessor subsystem 260. As discussed in greater detail below, the methodology of the present invention may be programmed and stored, as a set of program instructions for subsequent execution, in the processor arrangement 190 and its associated memory and other equivalent components.

In the preferred embodiment, the processor arrangement 190 is utilized, in conjunction with a stored set of program instructions and in response to any control signals entered by the user or received from the network 140, first, to convert the received protocol signal (from the wireline network interface 210) both to a baseband output video signal (to be modulated by the RF modulator/demodulator 205 and transmitted to a video display 225) and to an output audio signal (transmitted to the audio 220, or combined with the baseband output video signal and modulated and transmitted to the video display 225, or both); and second, to convert both a baseband input video signal (the demodulated input video signal having originated from the camera interface 235) and an input audio signal (from the audio 220 or combined with the baseband input video signal having originated from the video camera 230 and the camera interface 235), to the second protocol signal (to be modulated or formatted and transmitted by the wireline network interface 210 to the network 140). The functions of each of the components of the video access apparatus 110 are discussed in greater detail below with reference to FIGs. 3 - 10.

FIG. 3 is a high level block diagram illustrating a second embodiment of a video access apparatus, namely, video access apparatus 150, and illustrating a second embodiment of a video conferencing system 300, in accordance with the invention disclosed in the fifth related application. The second apparatus embodiment, namely, the video access apparatus 150 illustrated in FIG. 3, is the preferred apparatus embodiment of the invention, and is in all other respects equivalent to and may be utilized in a manner identical to the first embodiment, video access apparatus 110, illustrated in FIGs. 1 and 2. Similarly, the second embodiment of the video conferencing system, video conferencing system 300, is also the preferred system

embodiment, and is in all other respects equivalent to and may be utilized in a manner identical to the first embodiment, video conferencing system 200, illustrated in FIG. 2.

As illustrated in FIG. 3, the video access apparatus 150 includes a microprocessor subsystem 260 and an audio/video compression and decompression subsystem 265, which form the processor arrangement 190 discussed above with reference to FIG. 2. The video access apparatus 150 also includes an ISDN interface 245 and a telephony interface 250 (which individually or in combination equivalently function as the wireline network interface 210 illustrated in FIG. 2), a user audio interface 255 (which equivalently functions as the user interface 215 illustrated in FIG. 2); and an RF modulator 270 and RF demodulator 275 (which together equivalently function as the RF modulator/demodulator 205 illustrated in FIG. 2). In this preferred embodiment, the first communication channel 103 includes an ISDN or other digital line 105, coupleable to the ISDN interface 245, and a telephony (POTS) line 107, coupleable to the telephony interface 250. Depending upon the desired embodiment, discussed below with reference to FIG. 4, both the ISDN interface 245 (and corresponding digital line 105) and the telephony interface 250 (and corresponding telephony line 107) do not need to be included, as one or the other is sufficient. For example, a user or subscriber who does not desire an ISDN connection may choose an implementation of the video access apparatus 150 having only a telephony interface 250 (and corresponding telephony line 107), without an additional ISDN interface 245 (and corresponding digital line 105). The preferred embodiment of the video access apparatus 150 illustrated in FIG. 3 also includes a line or connector 115 for connection to a television antenna or to cable television for input of a television broadcast, cable

television or other video; a filter 285; and a directional coupler 290. The functions of each of these components is explained in greater detail below.

Also as illustrated in FIG. 3, the second embodiment of a video conferencing system 300 includes (as an audio interface) one or more telephones 295_1 through 295_n (individually and collectively referred to as telephone(s) 295, and which telephones 295 equivalently function as the audio 220 illustrated in FIG. 2); the video access apparatus 150; a video camera 230; a camera interface 235 (which also may be combined or incorporated within the video camera 230); one or more televisions 240_1 though 240_n (which are individually and collectively referred to as television(s) 240, and which equivalently function as the video displays 225 illustrated in FIG. 2); and a second communications channel 227 which, as mentioned above, is preferably a coaxial cable in the user (or subscriber) premises.

Referring to FIG. 3, the video access apparatus 150 provides both telephony (POTS) and audio/video conferencing service using common household appliances for interaction with the user (or subscriber) in the video conferencing system 300, such as telephones 295_1 through 295_n for entry of control signals and for audio input and output; video camera 230 for video input (such as a video camcorder); and television(s) 240 for video output (as or in lieu of video displays 225). When providing POTS service, the video access apparatus 150 interfaces with the typical, existing twisted-pair cabling 294 in the user (or subscriber) premises so that any telephone in the user premises, such as telephones 295_1 through 295_n, may be used. The video access apparatus 150 also provides line current and traditional "BORSHT" functions for typical (POTS) telephone service, as explained in greater detail below.

When providing video conferencing service, any of the plurality of telephones 295_1 through 295_n (individually and collectively referred to as telephone(s) 295) may be used for call (conference) establishment or set up and for audio input and output. The radio frequency output video signal (from the video access apparatus 150) may be displayed on any of the televisions 240 connected to the second communication channel 227 (such as a CATV coaxial cable) within the user premises, using any channel (when not connected to cable TV) or using any vacant channel within the CATV downstream frequency band (for example, channel 3 or 4). The radio frequency output video signal is originally received via the first communication channel 103 from the network 140 in a modulated or formatted digital form, such as digital data modulated and encoded utilizing one or more protocols such as H. 32x, and Q.x or V.x, which may be referred to as a received or first protocol signal. The first protocol signal is received over the first communication channel 103 of the audio/video network 100, having been transmitted via, for example, the network 140, from another, second user premises. The first protocol signal, typically consisting of encoded/modulated and compressed digital data, is received by the video access apparatus 150, which decodes/demodulates and decompresses the data and converts it to a baseband output video signal, such as an NTSC/PAL composite video signal (NTSC being a video format typically utilized in North America and Japan, with PAL being a video format typically utilized in Europe). Other video formats may also be used, such as SECAM (typically used in France) or HDTV (high definition television formats). This baseband output video signal (on line 271) is then RF modulated (using RF modulator 270) onto an available video RF carrier and injected into the second communication channel 227 (e.g., coaxial cable) at the user premises using a

directional coupler 290 (preferably 4 port). The radio frequency output video signal is then sent to all television receivers, such as televisions 240, within the user premises, such as a home or office. The directional coupler 290 is used in the preferred embodiment to provide directional signal injection while providing isolation with any connected CATV network (which may be coupled via line 115).

The video signal originating in the user premises and to be transmitted via the network 140 to another, second user premises (or any other location), originates from a video camera (or camcorder) 230 that produces a video signal, such as an NTSC/PAL composite video signal, which is also preferably modulated on channel 3 or 4 (61.25 or 67.25 MHz). This RF video signal from the video camera 230 is connected or coupled to a camera interface 235, which utilizes an offset mixer to shift the RF video signal (typically on a 61.25 or 67.25 MHz carrier) up to a spectrum higher than typical CATV frequencies, such as the 1.2 GHz or 900 MHz bands, to avoid interfering with the radio frequency output video signals or other CATV downstream channels. When the video access apparatus is not connected to CATV, such offset mixing may be unnecessary and the camera interface 235 may be omitted from the system 300, provided interference with the downstream radio frequency output video signals may be avoided (for example, utilizing downstream transmission on channel 9 and upstream (input) transmission on channel 3 or 4). For those video cameras 230 which may not include a modulator to shift the NTSC/PAL composite video signal to channel 3 or 4, such modulation may be incorporated into the camera interface 235; conversely, the functions of the camera interface 235 may also be incorporated directly into the video camera 230. The shifted (offset mixed) video signal from the camera interface 235 (or unshifted video signal directly from the camera 230, if CATV or other downstream interference is

not an issue), referred to herein as a radio frequency input video signal, is then injected into the same second communication channel 227 (also connected to the televisions 240) and transmitted to the video access apparatus 150. The video access apparatus 150 receives the radio frequency input video signal via the directional coupler (preferably at 1.2 GHz or 900 MHz) and demodulates the signal to baseband using RF demodulator 275, to form a baseband input video signal (on line 272). The baseband input video signal is then converted to digital form and compressed, to form a second protocol signal, such as a H.32x encoded video signal, and is transmitted (to form a transmitted protocol signal, which preferably also has further encoding and or modulation, such as a further Q.x or V.x encoded signal) over the audio/video network 100 via first communication channel 103. In the preferred embodiment, by using a vacant video channel at 1.2 GHz or 900 MHz, interference with any applicable downstream and upstream video, television, or CATV services tends to be avoided. The 1.2 GHz or 900 MHz signal is also filtered out of the feed-through cable or link 287 by a low pass filter 285, so that the signal is highly attenuated before it may leave the video access apparatus 150 through any cable which may be attached via line 115.

While the primary function of the video access apparatus 110 (or 150) and the video conferencing system 200 (or 300) is to provide full-duplex video communications, other secondary functions are also available in the preferred embodiment. For example, one such secondary function is a "loop back function" which allows the user to view the video from the video camera 230 on the screen of a television 240 or video display 225, such that the RF input video signal is demodulated (from 1.2 GHz or 900 MHz), remodulated onto a video RF carrier (that is tunable or receivable by the

televisions 240), and utilized for an RF output video signal. Such a loop back feature is especially valuable for surveillance, such as for home security or for baby monitoring. Also, a picture-in-picture (or multiple window) function may be provided, in which a user may view a small window of the video from video camera 230 along with the received video from another location, for example, to provide baby monitoring within the small window while simultaneously watching a movie or video received from a CATV network, or to provide a self-view for viewer feedback concerning the positioning of the viewer's own video camera 230.

In addition, the video access apparatus 110 (or 150) may be frequency agile, such that video conferencing may occur on any channel. While video conferencing on typically empty television or cable channels such as channels 3 or 4 may be preferable, in accordance with the present invention, video conferencing on additional channels is also feasible. For example, an existing video channel may be blanked out or eliminated, utilizing a notch filter, for any length of time, and the various input and output video signals inserted or overlaid into the now empty (filtered or muted) channel. Such frequency agility and injection of an audio/video signal, in the presence of existing programming, is one of many truly unique features of the present invention.

FIG. 4A is a block diagram illustrating a wireline network interface 210 of the preferred apparatus embodiment in accordance with the invention disclosed in the fifth related application. As indicated above, the wireline network interface 210 preferably is comprised of both an ISDN (digital) interface 245 and a telephony (or analog) interface 250, although either alone (digital or analog interface) is sufficient. As discussed in greater detail below, the first and second protocol signals, preferably encoded utilizing H.32x and

further encoded/modulated utilizing either Q.x or V.x protocols, are transported to and from the network 140 through one or both of these interfaces 245 and 250. Referring to FIG. 4A, utilizing an ISDN (digital) interface 245, connection to an ISDN or other digital network via line 105 is made through a jack 305 which may be, as discussed in greater detail below with reference to FIGs. 4B and 4C, for example, an RJ45 jack or an RJ11 jack, depending upon the service provided by the digital network. Coupled to the jack 305 is an isolation transformer circuit 310, which is further coupled to an ISDN transceiver 315 (which, as discussed below, may be either an S/T transceiver 315_a or a U transceiver 315_b). The ISDN transceiver 315, in turn, is coupled to the microprocessor subsystem 260 via a synchronous serial interface portion of bus 261.

FIG. 4B is a block diagram illustrating an ISDN S/T interface 245_a for use with pre-existing ISDN service. For example, a digital network service provider may typically bring a twisted pair line to the outside of a subscriber's premises, and install an ISDN interface. As a consequence, when there is a pre-existing ISDN NT1 interface, such as interface 306 (having an NT1 function for two to four wire conversion), appropriate connection to the existing NT1 interface should be made utilizing an ISDN S/T interface 245_a. As a consequence, as illustrated in FIG. 4B, the jack 305 is implemented as an RJ45 jack 305_a, the isolation transformer circuit 310 is implemented as an S/T dual isolation transformer 310_a, and the ISDN transceiver 315 is implemented as an ISDN S/T transceiver 315_a (such as a Motorola MC145574 integrated circuit).

FIG. 4C is a block diagram illustrating an ISDN U interface 245_b for use when there is no pre-existing ISDN service (having an installed NT1 interface). In this implementation, the jack 305 is implemented as an RJ11 jack

305b, the isolation transformer circuit 310 is implemented as a U isolation transformer 310b, and the ISDN transceiver 315 is implemented as an ISDN U transceiver 315b which also performs a NT1 function (such as a Motorola MC145572 integrated circuit).

Referring to FIG. 4A, for digital service, the ISDN interface 245 consists of an ISDN transceiver 315, such as the Motorola MC145574 or MC145572, and an isolation transformer circuit 310, which provide the layer one interface for the transportation of two 64 kbps B channels and one 16 kbps D channel between the network 140 termination (jack 305) and the microprocessor subsystem 260, preferably performing certain portions of the ISDN protocols, namely, Q.910 physical layer and Q.921 LAPD data link protocols. The ISDN transceiver 315 provides the modulation/line transmit and demodulation/line receive functions, as well as activation, deactivation, error monitoring, framing, and bit and octet timing. The ISDN transceiver 315 interfaces with the microprocessor subsystem 260 over a synchronous serial interface (SSI) portion of the bus 261. As discussed in greater detail below, the microprocessor subsystem 260 performs the Q.931 message signaling ISDN protocol and provides overall control of all subsystems within a video access apparatus 110 or 150, while the audio/video compression and decompression subsystem 265 performs the H.32x protocols.

Continuing to refer to FIG. 4A, for analog service, the telephony (or analog) interface 250 performs analog modem functions, operating, for example, as a V.34 or V.34bis modem. Connection to an analog network, via a telephony (POTS) line 107, is made via a jack 320, which is typically an RJ11 jack. Connected to the jack 320 is a dial (or data) access arrangement (DAA) 325, which receives an analog signal transmitted on the analog telephony line 107. DAAs are known in the prior art and may be made of a variety of

discrete components, including analog multiplexers, resistors, capacitors, and operational amplifiers, or may be embodied in whole or part as an integrated circuit, such as a Cermetek CH1837, and performs such functions as impedance matching, power level adjustment, isolation, surge voltage protection, and ring detection functions. Connected to the DAA 325 is a codec (coder-decoder) 330, such as a Motorola MC145500 integrated circuit (or, equivalently, an analog-digital (A/D) converter) which converts an analog signal received from the line 107 to sampled, digital form, and converts sampled, digital information to analog form for transmission over the line 107. The codec 330 is also referred to as a network codec 330, to distinguish it from a second codec, the audio codec 410, utilized in the user audio interface 255. The network codec 330 interfaces with a voice digital signal processor (DSP) 415 (of the user audio interface 255), also over a synchronous serial interface (SSI) portion of the bus 261. The network codec 330 performs V.x functions when in video mode, and voice functions when in telephony mode, as discussed in greater detail below. When utilized in this analog modem role (V.x functions), the voice DSP 415 operates in conjunction with the video processing DSP 365 (of the audio/video compression and decompression subsystem 265) utilizing a set of modem program instructions under the control of the microprocessor subsystem 260. The audio/video compression and decompression subsystem 265 also performs H.32x compression and decompression of the various input and output audio and video signals. This telephony interface 250 is used in the preferred embodiment for V.x modem functions during a video telephony call, and analog audio functions during a typical voice (POTS) call.

FIG. 5 is a block diagram illustrating a microprocessor subsystem 260 of the preferred apparatus embodiment in accordance with the invention disclosed in the fifth related

application. The microprocessor subsystem 260 consists of a microprocessor 350 or other processing unit, such as the Motorola MC68LC302, and memory 360, which includes random access memory (RAM) and read-only memory (ROM), and in the preferred embodiment, also includes flash programmable memory (such as flash EPROM or E^2PROM), with communication provided over the bus 261 to the ISDN interface 245, the user audio interface 255 (and voice DSP 415), and the audio/video compression and decompression subsystem 265. The read only memory portion of memory 360 also utilizes flash programmable memory, such that the memory contents may be downloaded over the audio/video network 100. As a consequence, different versions of operating software (program instructions), such as upgrades, may implemented without modifications to the video access apparatus 150 and without user intervention.

Continuing to refer to FIG. 5, the microprocessor subsystem 260 provides device control and configuration, call processing, and is also used to implement an ISDN protocol stack when required for video calls, such as Q.931 message signaling. Because the microprocessor subsystem interfaces with the ISDN interface 245 and the telephony interface 250 (via the voice DSP 415), a high speed data link may be established between the network 140 and the audio/video compression and decompression subsystem 265 using the microprocessor subsystem 260 as the data exchange and protocol conversion device. User audio, in the form of a pulse code modulated (PCM) data stream, may also be routed through the microprocessor 350 to the audio/video compression and decompression subsystem 265 from the voice DSP 415 of the user audio interface 255.

FIG. 6 is a block diagram illustrating an audio/video compression and decompression subsystem 265 of the preferred apparatus embodiment in accordance with the

invention disclosed in the fifth related application. The audio/video compression and decompression subsystem 265 performs video compression of the baseband input video signal (originating from the video camera 230 and camera interface 235) and audio compression of the input audio signal (from the user audio interface 255), and decompression of the audio and the video data of the received, first protocol signal (the first protocol previously having been decoded and/or demodulated) for subsequent display on the television(s) 240, all preferably utilizing the H.32x family of protocols. The audio/video compression and decompression subsystem 265 includes a video processing digital signal processor (DSP) 365, a red-green-blue digital to analog converter 370, a red-green-blue analog to digital converter 390, an encoder 375, and an audio/video input processor 380. The video processing DSP (or video processing DSP subsystem) 365 is a high-speed programmable DSP (or DSP arrangement or subsystem, such as a Motorola DSP56303 with associated support components, including memory and a hardware acceleration ASIC (discussed below), utilized to implement different video and audio compression and decompression algorithms, depending on the transmission rate and/or video conferencing standard at the remote end (i.e., the other premises with which the video access apparatus is communicating). The program code for the video processing DSP 365 may also be downloaded from the microprocessor subsystem memory 360, which may also be downloaded by a service provider through the audio/video network 100. As a consequence, video functionality of the video access apparatus 150, including new algorithms, may be changed or upgraded on-the-fly, also without any hardware changes and without user intervention.

Continuing to refer to FIG. 6, compressed audio/video data received from the network 140 (as, for example, H.32x

encoded protocol signals), via the ISDN interface 245 and the microprocessor subsystem 260, or via the telephony interface 250 with the voice DSP 415 and the microprocessor subsystem 260, is transferred to the video processing DSP 365 where it is decompressed, with video also converted to red-green-blue ("RGB") digital video signals, and with decompressed audio transferred to the user audio interface 255 (or modulated for transmission to the televisions 240). The RGB digital video signals are then converted to RGB analog signals, by the RGB digital to analog ("D/A") converter 370, such as the Motorola MC44200. The analog RGB signals, along with a composite synchronization signal, are then applied to an encoder 375, preferably an NTSC/PAL encoder such as a Motorola MC13077, resulting in an NTSC/PAL composite video signal, which may also be referred to as a baseband output video signal. The NTSC/PAL composite video signal is then transferred to the RF modulator 275 for upconversion to a radio frequency (to form the radio frequency output video signal), followed by transmission on the second communications channel 227 and display on a television 240.

For subsequent transmission over the network 140 of an input video signal (originating from the video camera 230 and the camera interface 235), a baseband input video signal, such as an NTSC/PAL composite video camera or camcorder signal, is received from the RF demodulator 270. The baseband input video signal is transferred to an audio/video input processor 380, such as a Motorola MC44011, which converts the baseband input video signal to analog RGB signals, while also providing a genlocked sampling clock for subsequent digitizing of the video signals. These input analog RGB signals are then converted to digital RGB signals by a RGB analog to digital converter 390, such as the Motorola MC44250, and transferred to the video processing DSP 365. The video processing DSP 365 compresses the digital RGB

signals and audio data (from the user audio interface 255), preferably utilizing an H.32x protocol, and transfers the resulting data stream to the microprocessor subsystem 260 for additional analog or digital processing. It should be noted that as part of the H.32x protocol, audio information originating from the user audio interface 255 or from the video camera 230 (and camera interface 235) is compressed and combined with compressed video data before transmission to the network 140 via the wireline network interface 210. For subsequent digital transmission, the microprocessor subsystem 260 encodes the compressed audio/video data utilizing the Q.931 ISDN message signaling protocol, and transfers the processed data to the ISDN interface 245, for additional ISDN protocol processing and transmission over the first communication channel 103. For subsequent analog transmission, the microprocessor subsystem 260, the voice DSP 415 (of the user audio interface 255) and the video processing DSP 365 encode the compressed audio/video data utilizing analog protocols such as the V.x series of protocols, and transfer the processed data to the telephony interface 250, for additional V.x protocol processing and transmission over the first communication channel 103. In the preferred embodiment, the audio/video compression and decompression subsystem 265 may also include additional random access memory for use by the video processing DSP 365 for partial or full storage of pixel data of an input/output video frame. Also in the preferred embodiment, a hardware acceleration ASIC is used to assist the video processing DSP 365 in processing speed intensive tasks, such as discrete cosine transforms associated with the compression and decompression processes.

FIG. 7 is a block diagram illustrating a user audio interface 255 of the preferred apparatus embodiment in accordance with the invention disclosed in the fifth related

application. The user audio interface 255 is designed to interface with standard household telephone sets, including wireless devices and speaker phones, such as telephones 295_1 through 295_n. The user audio interface 255 is intended to support both audio POTS calls and video calls. In the preferred embodiment, POTS calls are processed in a "transparent" mode, such that placing and receiving telephone calls occur as if no video call functions were present. Also in the preferred embodiment, video calls are processed as an exception, requiring a designated or predetermined dialing sequence entered by the user to invoke a video call.

Referring to FIG. 7, a SLIC (Subscriber Loop Interface Circuit) 400 provides "BORSHT" functions for telephone service within the user premises, such as that normally provided by a network central office, including DC (direct current) power for the telephone (Battery); Overvoltage protection; Ring trip detection and facilitation of ringing insertion; Supervision features such as hook status and dial pulsing; Hybrid features such as two-wire differential to four-wire single-ended conversions and suppression of longitudinal signals at the two-wire input; and Testing. The SLIC 400 communicates with the telephones 295_1 through 295_n through an ordinary telephone line, such as twisted pair cabling 294, which has tip and ring lines. The ring generator 405 provides high-voltage AC (alternating current) signals to ring the telephones 295_1 through 295_n. Connected to the SLIC 400, the audio codec 410 provides analog-to-digital conversion for voice digitizing of the input (voice) audio signal originating from the microphone portion of one or more of the telephones 295_1 through 295_n, to form an input (PCM) digital voice data stream or signal, and digital-to-analog conversion for voice recovery from an output (PCM) digital voice data stream or signal (to create the output audio

signal to the speaker portion of the telephones 295_1 through 295_n), and well as band limiting and signal restoration for PCM systems. The output and input (PCM) digital voice data streams connect directly to the voice processing DSP 415. The voice processing DSP 415, such as a Motorola DSP56303, contains program memory and data memory to perform signal processing functions such as DTMF/dial pulse detection and generation, analog modem functions, call progress tone (dial tone, busy tone) generation, PCM-to-linear and linear-to-PCM conversion, and speech prompt playback. As indicated above, the voice processing DSP 415 also provides modem functions, such as V.x modem functions, to additionally support POTS or other analog-based video calls. The voice processing DSP 415 interfaces with the microprocessor subsystem 260 and network codec 330 over the bus 261. The memory 420 (connected to the voice processing DSP 415), in the preferred embodiment, includes high density read only memory (referred to as speech ROM) containing PCM encoded (or compressed) speech segments used for interaction with the user, such as in prompting the user for keypad DTMF or dial pulse entry when in the video calling mode. In addition, optional speech random access memory may be used for user voice storage functions, and electrically alterable, programmable non-volatile (flash) memory for storage of programs (and updates) or algorithms.

The user audio interface 255, in the preferred embodiment, operates in one of two modes, first, for telephony (POTS), and second, for video conferencing (calling). The telephony (POTS) mode is user transparent, as a default mode which is entered whenever the user goes off hook. As discussed in greater detail below, the video conferencing mode is entered as an exception, through the user entering (dialing) a specific, predetermined sequence which, in the preferred embodiment, is not recognized as a

telephony sequence. In the telephony (POTS) mode, the voice processing DSP 415 generates the customary "dial" tone when the user telephone (of the telephones 295_1 through 295_n) goes off hook. The user then enters the dialing sequence via the keypad of a telephone 295, just as in known or customary telephone dialing. The voice processing DSP 415 decodes the dialing digits and stores them in a calling memory buffer of memory 420. Upon decoding the first two digits entered (which are not the first two digits of the specific predetermined video call sequence), the voice processing DSP 415 recognizes that the requested call is not a video call and, as a consequence, signals the microprocessor subsystem 260 to initiate a POTS call through the audio/video network 100 using the telephony (analog) interface 250. When the call is granted (by the network 140) and the audio link with the local digital or analog switch 135 is established, the voice processing DSP 415 forwards the stored digits to the local digital or analog switch 135 and connects the audio paths between the user's telephone(s) and the network 140. From this point on, the voice processing DSP 415 will not decode any dialed digits and will simply pass through the input and output PCM digital voice data stream, until the user's telephone goes on hook and the call is terminated.

Alternatively for a telephony session, the audio/user interface 255 may create or maintain a connection to a central office of a network 140, to provide transparency for telephony. Once the entry of the specific predetermined sequence for video mode is detected, the audio/user interface 255 breaks or terminates the central office connection, and enters video mode, under local control of the video access apparatus 150 (or 110).

As indicated above, the user initiates the video conferencing mode as an exception to the normal telephony mode, by entering a specific predetermined sequence which is

recognized by the voice processing DSP 415 as a non-telephony sequence and, additionally in the preferred embodiment, as the predetermined sequence specific to the video mode. This methodology is also discussed below with reference to the flow chart of FIG. 12. For the video conference mode of the preferred embodiment, the first two digits of the specific, predetermined sequence are unique and specifically unused in a standard POTS call, such as "* *", and as a consequence, may specifically signal the audio voice processing DSP 415 to enter the video call mode. Alternatively, other specific, predetermined sequences could be programmed by the user for recognition as a video conference mode by the voice processing DSP 415. Immediately after decoding the two special digits or other specific predetermined sequence, the voice processing DSP 415 generates or plays a speech prompt sequence, such as "Please select a call option or press the '#' key for help", which is stored in the speech ROM portion of memory 420. The action taken by the voice processing DSP 415 will then depend upon the sequence entered or key pressed by the user following the initial prompt. For example, if the '#' key is pressed, the user may hear a menu of commands such as, for example, the following:

- "To place a Directory call, press *"
- "To update the call Directory, press 2"
- "To place a manual video call, press 3"
- "To mute the camera, press 4"
- "To view the camera on your television, press 5"
- "To hear this menu again, press #'

Thus, in the preferred embodiment, an automated and user friendly prompting sequence is used to guide the user through placing a video conference call. Once the entry is

complete, the information is then passed from the voice processing DSP 415 to the microprocessor subsystem 260, which will then attempt to connect the call through the network 140. If successful, the audio paths (input and output audio signals) will be connected through to the telephones 295_1 through 295_n, the output video path will be connected through to the televisions 240_1 through 240_n (or other video displays 225), and the input video path will be connected from the camera interface 235 (originating from the video camera 230). Alternatively, under user or other control, the output audio path may also be connected to a television 240, for broadcast over the speakers within the television(s) 240, and the input audio path may also originate from a microphone within the video camera 230 and be connected via the camera interface 235. This alternate path may be particularly useful when the user desires to video tape the video conference, for example, utilizing an ordinary VCR coupled to the television 240. The video call terminates when the telephone goes on hook, or another control signal is entered via the user interface 215 or user audio interface 255.

It should be noted that in the preferred embodiment, a simple directory feature may be used to simplify the video calling process. For example, after the user goes off hook and presses the '*' key three times followed by a single digit '1', '2'...'9', a call automatically may be placed using a sequence of numbers stored in the directory for that digit. This feature may be necessary or desirable under a variety of circumstances, for example, when an ISDN call may require the entry of two separate 10-digit numbers to connect the call through the network 140. Also as an option in the preferred embodiment, a more sophisticated system may store a simple name tag or other alphanumeric entry associated with the directory entry, created by the user, and played back to the user by the voice processing DSP 415. For example, a prompt in

response to making a directory call may be: "To call 'grandma', press 1"; "To call 'mother', press 2"; "To call 'work', press 3"; in which the speech segments "grandma", "mother", and "work" are spoken by the user, recorded and stored in memory 420. More sophisticated systems may include speaker/voice recognition techniques, to recognize the user selection, eliminating the need to press any keys on a telephone keypad or other manual entry of information into the user interface 215 or user audio interface 255. It should also be noted that video call control functions, such as camera muting, unmuting, and local playback (loop back), also may be selected with the same user interface. Other sophisticated systems may also include use of the video display 225 or television 240 for on-screen visual display of a menu of options, with corresponding entry of user control signals, such as call control and placement information, occurring in a variety of ways, such as through the keypad of the telephones 295, through a infrared remote control link with the video access apparatus 150 (or 110), or through the input video path via the second communication channel 227. In this manner, the keypad or remote control link, coupled with the video display, may effectively form a distributed graphical user interface for call control. These various methods of user prompting, on-screen display, and user feedback are especially useful to guide the user through the process of placing a video call, and help to make the audio video conferencing system 300 (or 200) especially user-friendly. In addition, these various methods also illustrate the "tri-ality" of the use of a telephone 295 in the preferred embodiment, for telephony, for audio input and output, and for call control.

FIG. 8 is a block diagram illustrating an RF modulator 270 of the preferred apparatus embodiment in accordance with the invention disclosed in the fifth related application. The RF modulator 270 converts the baseband output video

signal from the audio/video compression and decompression subsystem 265, such as an NTSC/PAL composite video signal, to a radio frequency output video signal, such as an amplitude modulated vestigial sideband RF signal, which may be viewed via the receiver of the user's television 240, for example, when tuned to channel 3 or 4. The RF modulator 270 may be implemented in a variety of ways, including through use of a video modulator 425, such as a Motorola MC1373, followed by a gain stage (amplifier) 430, utilized in the preferred embodiment to overcome losses from the directional coupler 290 which feeds the RF output video signal into the second communication channel 227, such as a coaxial cable system in the user premises. A switchable notch filter may also be used to remove current programming from a particular channel (RF video carrier), while inserting the radio frequency output video signal into the second communication channel 227.

FIG. 9 is a block diagram illustrating an RF demodulator 275 of the preferred apparatus embodiment in accordance with the invention disclosed in the fifth related application. In the preferred embodiment, the RF demodulator 275 is a full heterodyne receiver tuned to a specific channel in the 900 MHz band or 1.2 GHz band, to receive the radio frequency input video signal from the camera interface 235 (originating from the video camera 230). The radio frequency input video signal, fed into the RF demodulator 275 from the directional coupler 290, is bandpass filtered (at either 900 MHz or 1.2 GHz) in pre-filter 435, then mixed down to an intermediate frequency (IF) of, for example, 45 MHz, using the mixer 440 and a fixed reference oscillator 445. The signal is then surface acoustic wave (SAW) filtered by the SAW filter 450, or otherwise bandpass filtered, and transferred to a (color) TV IF subsystem 460, such as a Motorola MC44301, which provides amplification, AM

detection (demodulation) and automatic fine tuning, resulting in a baseband input video signal (baseband composite input video signal). This baseband input video signal is then transferred to the audio/video compression and decompression subsystem 265 for further processing as discussed above.

FIG. 10 is a block diagram illustrating a camera interface 235 of the preferred apparatus embodiment in accordance with the invention disclosed in the fifth related application. The camera interface 235 is used in conjunction with a video camera (or camcorder) 230 that outputs its signal as an RF video carrier on channel 3 or 4 (61.25 or 67.25 MHz), and is used to upconvert the video carrier to an RF carrier at 900 MHz or 1.2 GHz without intervening demodulation and modulation of the video signal. As mentioned above, the camera interface 235 may be omitted when the video access apparatus 150 (or 110) is not connected to CATV services and, in which case, the video camera 230 may be directly connected to the second communication channel 227 (provided that interference with the RF output video signal may be avoided, for example, by having the RF input video signal from the video camera 230 on a different channel than the RF output video signal from video access apparatus 150). As illustrated in FIG. 10, the input video signal from the video camera 230 is mixed up to the required output frequency using an offset mixer 465, a fixed reference oscillator 470, and a bandpass filter 475. Not illustrated in FIG. 10, if additional input video signals are desired from, for example, additional video cameras, the input video signals may also be multiplexed. This feature may be desirable, for example, when the system is to be used for surveillance of multiple points or locations, or when the user desires to transmit additional windows or screens within screens.

Alternatively, as mentioned above, the camera interface 235 may be directly incorporated within the video camera 230. In addition, for those video cameras producing a NTSC/PAL composite video signal (rather than an RF video carrier on channel 3 or 4), an additional stage may be added within the camera interface 235 to modulate the NTSC/PAL composite video signal to an RF video carrier prior to offset mixing by offset mixer 465, or in lieu of offset mixing, directly modulating the NTSC/PAL composite video signal to 900 MHz or 1.2 GHz to form the RF input video signal.

FIG. 11 is a flow diagram illustrating the method of the preferred embodiment of the invention disclosed in the fifth related application. As illustrated in FIG. 11, the method begins, start step 500, with receiving a first protocol signal, such as a Q.x or V.x encoded/modulated H.32x audio/video signal, to form a received protocol signal, step 505. In the preferred embodiment, step 505 is performed in the wireline network interface 210. Next, in step 515, the received protocol signal is converted to a baseband output video signal and an output audio signal. In the preferred embodiment, step 515 is performed by the processor arrangement 190, or more particularly by the microprocessor subsystem 260 (and possibly voice DSP 415) and the audio/video compression and decompression subsystem 265. In the preferred embodiment utilizing audio 220 or telephones 295 for audio output and input, an important feature of the present invention is the independence of the output audio signal from the output video signal. In the event that a television 240 or other video display 225 is also to be used for audio output, the output audio signal may be combined with the baseband output video signal (rather than separating out the audio portion and separately routing it to audio 220 or telephones 295_1 through 295_n). Next, in step 525, the baseband output video signal (and possibly output audio signal as well) is modulated

to form a radio frequency output video (and audio) signal, also referred to as a composite output video signal, and in step 535, the RF output video (and audio) signal is transmitted. In the preferred embodiment, steps 525 and 535 are performed by the RF modulator/demodulator 205 or the RF modulator 270. In addition, the output audio signal may also be a combination of both near end and far end (remote) audio, resulting in near and far end combined audio available at the television receiver. This combination would allow both recording and monitoring of the audio/video information, from both the near and far ends.

Concurrently with steps 505, 515, 525 and 535 (involving receiving (at a local location) video conference information transmitted from another location, such as a remote location), in the preferred embodiment, steps 510, 520, 530 and 540 are also occurring (involving transmitting (from a local location) video conference information to another location, such as a remote location). In step 510, a radio frequency input video signal and an input audio signal are received. As indicated above, in the preferred embodiment, the input video signal and input audio signal are each independent of the other. In the preferred embodiment, the radio frequency input video signal from the camera interface 235 (or directly from the video camera 230) is received by the RF demodulator 275 or the RF modulator/demodulator 205, and an input audio signal is received by either the audio 220 and user interface 215, or telephones 295_1 through 295_n and user audio interface 255. Alternatively, the input audio signal may also be received by a microphone in the video camera 230 and included as part of the RF input video signal from the camera interface 235. Next, preferably in the RF demodulator 275 or the RF modulator/demodulator 205, in step 520 the RF input video (and possibly audio) signal is demodulated to form a baseband input video (and possibly audio) signal. In

step 530, the baseband input video signal and the input audio signal are converted to a second protocol signal, preferably by the processor arrangement 190, or more specifically by the audio/video compression and decompression subsystem 265, the microprocessor subsystem 260, and the voice DSP 415. In step 540, the second protocol signal is transmitted to form a transmitted protocol signal, preferably by the wireline network interface 210. Following steps 535 and 540, when the video conference has been terminated, step 545, such as by going on hook, the process may end, return step 550, and if the video conference has not been terminated in step 545, the method continues, returning to steps 505 and 510.

FIG. 12 is a flow diagram illustrating the telephony and video conference control methodology in accordance with the preferred embodiment of the invention disclosed in the fifth related application. FIG. 12 also illustrates the multiple roles of a telephone, such as telephones 295_1 through 295_n, in the system of the present invention, including providing telephony (POTS), providing video call control, and providing the audio portion of the video conference. Referring to FIG. 12, beginning with start step 600, a request for service is detected, step 605, such as going off hook or receiving an incoming alert signal. Next, in step 610, a user indication or alert is provided, such as a dial tone or an incoming ring signal, and signaling information is collected, such as DTMF digits of a phone number or "**". When a video conference has been requested in step 615, such as through entry of "**" or receipt of an incoming message from the network 140, then the method proceeds to step 635. When a video conference has not been requested in step 615, the method proceeds to request or set up a telephony call, such as generating DTMF tones and connecting an audio path between the user's telephone and the network 140, step 620, followed by entering the transparent telephony mode and

transmitting audio (typically PCM) data to the network 140, step 625. The audio data will have been PCM encoded, and will have been transformed into an appropriate digital or analog format (e.g., ISDN, POTS, etc.) by the wireline network interface 210 for transmission to the network 140. When the telephony call is terminated, step 630, the method may end, return step 660.

Continuing to refer to FIG. 12, when a video conference has been requested in step 615, the method proceeds to step 635 and initializes the video conference control system, such as playing an initial speech prompt as discussed above. Next, in step 640, the video input request type is collected and the corresponding requested service is performed, such as originating a video conference call using a directory, updating a video conference call directory, manually originating a video conference call, muting an input (audio or video), providing loop back (e.g., local self-view such as monitoring or other surveillance), playing help or error messages or menu options, or exiting the video conferencing control system. In step 645, a video conference call is requested or set up (such as for an incoming video call), and in step 650, the video conference mode is entered, with protocol encoded (e.g., H.32x and either Q.x or V.x protocols) audio and video data being transmitted to the network 140. When the video conference call is terminated in step 655, such as by going on hook, the method may end, return step 660.

A particularly innovative feature of the various apparatus and system embodiments of the present invention is the interoperability of both ISDN and POTS within the same apparatus and system. This ISDN/POTS interoperability is also related to the "tri-ality" of the use of a telephone 295 in the preferred embodiment for telephony (POTS), for audio input and output (for video, either ISDN or POTS), and for call control (for selecting either video or

telephony modes). As a consequence, when audio 220 or a telephone 295 may be being used for ISDN video conferencing, the method of the invention may include various modes for avoiding potential conflict with simultaneous POTS use. For example, during an ISDN video conference in which a telephone 295 is being utilized for call control and for audio input and output, the method provides for avoiding a POTS conflict, such as that which could occur if an incoming POTS call were received. One alternative for avoiding such conflict would consist of "busying out" the POTS line 107 when such an ISDN video conference is in progress. Another alternative would consist of providing POTS priority for the audio portion of the video conference, such as enabling a user to simultaneously receive the POTS audio while the video conference is occurring (or the video link maintained), for example, to provide for potentially exigent or emergency situations (such as emergency calls) which would typically occur via POTS lines. Other alternatives may include providing a POTS caller identification (caller ID) functionality, such that caller ID FSK modulated data could be displayed on a caller ID unit or on a video display 225 (or television 240), allowing the user to determine whether the video conference should or should not be terminated. Such an alternative may be implemented, for example, through a call waiting (flash hook) system, or by returning the POTS line to an on hook status followed by a ringing signal and going off hook. Similar conflict resolution schemes may be implemented for situations of an existing POTS call in progress followed by an incoming ISDN video call. In addition, a local, non-network flash system may also be implemented, allowing the user to toggle between a POTS call and a concurrent ISDN video call. Also as indicated above, such conflict resolution may also be implemented utilizing the combination of the keypad of a telephone 295

and video display 225 (or television 240) as a graphical user interface, for entry of user control signals and for selection of potentially competing calls.

Network configuration is yet another function which may be performed via a telephone 295 and user audio interface 255, especially utilizing menu options displayed utilizing an on screen display (on a television 220 or video display 225). For example, as disclosed in the fourth related application, automatic ISDN configuration capabilities, for example, for ISDN parameters such as switch type and SPID, may be implemented within the processor arrangement 190 and executed by the user via control functionality (as options entered by the user via the telephone 295 or other user interface 215). In addition, for POTS video conferencing capability, V.x or other modem configuration parameters (such as auto or manual answer) may also be configured as options entered by the user via the telephone 295 or other user interface 215.

The auto answer modem option also generates another possible area of conflict for POTS telephony versus POTS video conferencing, especially if a user is utilizing a telephone answering machine on the telephony (POTS) line 107. In the preferred embodiment, to determine whether an incoming POTS call is for telephony or video conferencing, in the preferred embodiment, a carrier (such as a V.34 carrier frequency) detector may be implemented, such that if a carrier is found, the video access apparatus 110 or 150 proceeds with V.x protocols (such as training), and if no carrier is detected, the video access apparatus 110 or 150 assumes a voice (telephony) call and allows the telephone 295 (or answering machine) to ring and answer the incoming call.

Similarly for ISDN telephony versus ISDN video conferencing, the video access apparatus 110 or 150 may detect an H.320 or other video protocol, and may provide a

distinctive alert to indicate an incoming video call. If the user then goes off hook, then the ISDN video call is connected, for example, using the Q.931 protocol. Correspondingly, if an answering machine goes off hook, the audio portion of the ISDN call may be passed through, allowing an audio message to be left during, for example, an H.320 video conference call.

FIG. 13 is a block diagram illustrating a third embodiment of a video access apparatus 750, and a third embodiment of a video conferencing system 705 utilizing a plurality of videophone apparatuses 700, in accordance with the present invention. Referring to FIG. 13, a plurality of videophone apparatuses 700_1, 700_2 through 700_n (individually and collectively referred to as videophone(s) 700 or videophone apparatus(es) 700), are connected to the video access apparatus 705 via a second communication channel 227, such as a coaxial cable within the user premises, and via a third communication channel 293, such as a POTS twisted pair cabling arrangement (which may be identical to line 294 illustrated in FIG. 3). The third embodiment of a video conferencing system 705 comprises one or more of the plurality of videophone apparatuses 700, coupled through the second communication channel 227 and the third communication channel 293 to the video access apparatus 750. Via the first communication channel 103, the video access apparatus 750 may also communicate with the network 140. (Given the plurality of communication channels 103, 227 and 293, for ease of description in the claims enumerated below, when used with reference to a videophone apparatus 700 which is coupleable to the two communication channels 227 and 293, the second communication channel 227 may be referred to as a first communication channel, with the third communication channel 293 referred to as a second communication channel).

Continuing to refer to FIG. 13, the video access apparatus 750 is very similar to the video access apparatus 150 discussed above with reference to FIGs. 3 through 9, and includes many of the same components having the same arrangement and operating in an identical fashion, including the wireline network interface 210 (preferably having both an ISDN interface 245 coupled to an ISDN line 105 and a telephony interface 250 coupled to a telephony (POTS) line 107, with lines 105 and 107 being part of or comprising the first communication channel 103); the user/audio interface 255; the microprocessor subsystem 260; the audio/video compression and decompression subsystem 265; the RF modulator 270; the filter 285; and the directional coupler 290. The video access apparatus 750 differs from the video access apparatus 150 insofar as the video access apparatus 750 has a RF carrier detector and switch 276, and has a plurality of RF demodulators 275, namely, RF demodulator 275_a, RF demodulator 275_b through RF demodulator 275_n (each of which is identical to the RF demodulator 275 illustrated in FIG. 9, and are individually and collectively referred to as RF demodulators 275). The video access apparatus 750 is utilized to multiplex or combine a plurality of input video signals from a plurality of videophones 700 (or 800); when such multiplexing or combining is unnecessary, any of the other video access apparatus embodiments (110 or 150) may also be utilized instead of the video access apparatus 750 in the videoconferencing system 750 illustrated in FIG. 13.

Continuing to refer to FIG. 13, each of the RF demodulators 275 are utilized to receive, on different channels or carriers (such as channels or carriers "a" through "n"), a radio frequency input video signal transmitted from a corresponding videophone apparatus 700 via the second communication channel 227, and to demodulate (downconvert) the received radio frequency input video

signal to a baseband input video signal. For example, RF demodulator 275_a, tuned to a radio frequency corresponding to channel "a" (e.g., tuned via oscillator 445 and mixer 440), may receive and convert to baseband a radio frequency input video signal transmitted on channel (or carrier) "a" from videophone apparatus 700_a, while RF demodulator 275_b, tuned to a radio frequency corresponding to channel "b", may receive and convert to baseband a radio frequency input video signal transmitted on channel (or carrier) "b" from videophone apparatus 700_b, and so on. In the preferred embodiment, when the video access apparatus is also coupled to a CATV cable via line 115, the various channels or carriers "a" through "n" are all at radio frequencies above the typical CATV spectrum, such as above 900 MHz, to avoid interference with downstream video, television, or other cable channels. For non-multiplexed input video signals, the video access apparatus may utilize the RF carrier detector and switch 276 to select which input video signal will be further transmitted to the network 140, potentially from among a plurality of input video signals (from each of the videophone apparatuses 700). In the event that more than one videophone apparatus 700 is transmitting a radio frequency input video signal to the video access apparatus 750, the RF carrier detector and switch 276 may, for example, select a particular channel (and corresponding videophone apparatus 700) on a contention basis, such as the first baseband input video signal to arrive at the RF carrier detector and switch 276, and transfer the baseband input video signal from that particular, selected channel to the audio/video compression and decompression subsystem 265, for further processing as discussed above with reference to FIGs. 3 through 9. In addition, via line or bus 277, the RF carrier detector and switch 276 also may be controlled by the microprocessor subsystem 260.

FIG. 14 is a detailed block diagram illustrating a first embodiment of a videophone apparatus 700 in accordance with the present invention. As illustrated in FIG. 14, the first embodiment of the videophone apparatus, videophone apparatus 700, incorporates corresponding features or elements of the first embodiment of the video conferencing system 200 (or, equivalently, of the second embodiment of the video conferencing system 300), which features or elements operate and function as discussed above. These incorporated, corresponding features or elements include features such as the video display 225 (or television 240), audio 220 (or telephone 295), the camera interface 235, and the video camera 230. More specifically, referring to FIG. 14, the videophone apparatus comprises: first, a video monitor 715, such as a television receiver and display (having the same functions as either the video display 225 or television 240); second, a camera 720 (having the same functions as the video camera 230); third, a camera interface 235 (as illustrated in FIG. 10); and fourth, a telephony module 710 (having the same functions as the audio 220 or telephone 295). As indicated above, a videophone 700 (of the plurality of videophones 700) may transmit a radio frequency input video signal (on a particular channel or carrier) to the video access apparatus 750, via the second communication channel 227, originating from the camera interface 235 and the camera 720. As discussed above, the camera 720 video output, preferably modulated onto channel 3 or 4, is offset mixed by the camera interface 235 to an assigned channel or carrier, preferably having a frequency above typical CATV frequencies, such as in the 900 MHz to 1.2 GHz spectrum. Each of the plurality of videophones 700 may receive, via the video monitor 715, a radio frequency output video signal from the video access apparatus 750 (via the second communication channel 227), originating from the network 140. An audio signal may also

be included within these video signals or, in the preferred embodiment, the audio signals (input and output) are transmitted and received independently via the telephony module 710 within the videophone apparatus 700, over the third communication channel 293. As discussed above concerning the telephones 295, in the preferred embodiment, video control is also provided via the telephony module 710, such as through user input (DTMF dialing). A videophone 700 may also include a directional coupler or splitter (not illustrated in FIG. 13) for connection or coupling of the radio frequency video signals to the second communication channel 227.

As indicated above, each of the various videophone apparatuses 700 may transmit a radio frequency input video signal to the video access apparatus 750, on an assigned channel or carrier of a plurality of channels or carriers, such as channel "a", channel "b", and so on. The assigned channel (or radio frequency carrier) may be assigned or otherwise determined through frequency selection via the oscillator 470 and mixer 465 of the camera interface illustrated in FIG. 10. Such a channel assignment may be set manually or downloaded at any time from the video access apparatus 750 via the second communication channel 227.

In the preferred embodiment, any of the videophone apparatuses 700 may receive a radio frequency output video signal (via the second communication channel 227), and may receive an output audio signal (via the third communication channel 293), both broadcast from the video access apparatus 750. As a consequence, any and all videophones 700 of the third embodiment of the video conferencing system 705 may display the output video and audio signals simultaneously. With regard to the radio frequency input video signal and input audio signal (transmitted to the video access apparatus 750), the input video signal originating from a particular

camera 720 in a videophone 700 may be selected, as discussed above, through channel selection within the video access apparatus 750, utilizing a selecting means such as the RF carrier detector and switch 276. In addition, the input video signal originating from a particular camera 720 in a videophone 700 also may be selected by the user, via entry of a control signal (of a plurality of control signals) (discussed in greater detail below with reference to FIG. 15), or two or more input video signals originating from more than one particular camera 720 may be selected and multiplexed to create a combined, multiple video signal, for example, utilizing a window or a split screen format (discussed in greater detail below with reference to FIG. 16).

FIG. 15 is a detailed block diagram illustrating a second embodiment of a videophone apparatus 800 in accordance with the present invention. As in the videophone apparatus 700 discussed above, the videophone apparatus 800 also includes a video monitor 715, a camera 720, a camera interface 235, and a telephony module 710. Like a videophone 700, a videophone 800 (of a plurality of videophones 800) also may transmit a radio frequency input video signal (on a particular channel or carrier) to the video access apparatus 750, via the second communication channel 227, originating from the camera interface 235 and the camera 720. As discussed above, the camera 720 video output, preferably modulated onto channel 3 or 4, is offset mixed by the camera interface 235 to an assigned channel or carrier, preferably having a frequency above typical CATV frequencies, such as in the 900 MHz to 1.2 GHz spectrum. Each of the plurality of videophones 800 may receive, via the video monitor 715, a radio frequency output video signal from the video access apparatus 750 (via the second communication channel 227), originating from the network 140. Also like the videophone 700, an audio signal may also be included within these video signals or, in the

preferred embodiment, the audio signals (input and output) are transmitted and received independently via the telephony module 710 within the videophone apparatus 800, over the third communication channel 293. As discussed above concerning the telephones 295, in the preferred embodiment, all of the various forms of video control are also provided via the telephony module 710, such as through user input (DTMF dialing) and on screen display.

As illustrated in FIG. 15, in the videophone apparatus 800, connected or coupled between the camera interface 235 and the second communication channel 227 is a selection switch 735, which is utilized to allow or prevent the input of a radio frequency input video signal, from the camera interface 235, into the second communication channel 227. The selection switch 735 is controlled through a detector 730, connected or coupled to the telephony module 710 and the third communication channel 293, which receives a control signal input by a user from any one of the telephony modules 710 of the plurality of videophones 800 coupled to the third communication channel 293. In the preferred embodiment, such control signaling is implemented utilizing DTMF dialing available in the telephony module. Alternatively, such control signaling may also be implemented utilizing other signaling or communications methods, such as frequency shift keying (FSK), or any of the other protocol or modulation methods discussed above. In addition to such signaling based or contention based control, as another alternative, such input control may be transparent to the user, for example, when the detector 730 and selection switch 735 are implemented utilizing speech detection, such as a voice operated switch (VOX), such that video is input automatically into the second communication channel 227 based upon whomever is speaking at that time.

Continuing to refer to FIG. 15, the videophone 800 which is to be the source of the input video signal (to the video access apparatus 750), of the plurality of videophones 800, is selected by the user through the telephony module 710, which is coupled via the third communication channel 293 to the user/audio interface 255 of the video access apparatus 750. In its simplest form, the detector 730 may simply detect which telephony module 710 or a particular videophone 800 (of the videophones 800) has gone off hook, and allows the selection switch 735 to allow or connect the input video signal from that off hook videophone 800. In the preferred embodiment, the user enters a predetermined DTMF sequence, such as a series of DTMF digits, to select the source videophone 800 which will provide the input video signal. In this implementation, all videophones 800 transmit a radio frequency input video signal on the same channel (RF carrier frequency), so that selection of a particular source videophone 800 avoids user contention and concomitant distortion which could result from simultaneous transmissions from multiple videophones 800, for example, when multiple videophones 800 have gone off hook. The DTMF sequence entered by the user is then utilized by the detector 730, via the selection switch 735, to enable or disable transmission of a radio frequency input video signal from a camera interface 235 (and camera 720) of any particular videophone 800. With this implementation, only a single RF demodulator 275 is required in the video access apparatus 750, as only one RF carrier is involved. The DTMF sequence entered by the user may also be utilized to automatically switch the video monitor 715 to the proper receive channel for reception of the radio frequency output video signal. In addition, when utilized with a camera interface 235 which may transmit on one or more assigned channels of a plurality of channels, for multiplexed radio frequency input video signals, the DTMF

sequence may also be utilized to select which videophones 800 will be sources for the radio frequency input video signals. Such multiplexing is discussed below with reference to FIG. 16.

FIG. 16 is a detailed block diagram illustrating a fourth embodiment of a video access apparatus 850, in accordance with the present invention. The video access apparatus 850 is utilized, in the preferred embodiment, for multiplexing or combining a plurality of input video signals, from a plurality of videophones 700 or 800, into one combined video signal for subsequent transmission over the first communication channel 103. For example, four input video signals, from four videophones 700 or 800, may be combined in a two by two (2 x 2) format (four windows) into one video signal, before compression and transmission through the network 140. To accomplish such multiplexing or combining, in addition to having the plurality of RF demodulators 275_a through 275_n of the video access apparatus 700 discussed above with reference to FIG. 13, the video access apparatus 800 also has an expanded audio/video compression and decompression subsystem 760. The expanded audio/video compression and decompression subsystem 760, illustrated in FIG. 17, performs the same functions as the audio/video compression and decompression subsystem 265 discussed above, but is able to perform these functions for a greater number of channels, rather than for a single channel.

FIG. 17 is a block diagram illustrating the expanded audio/video compression and decompression subsystem 760. The expanded audio/video compression and decompression subsystem 760 includes many of the same components as the audio/video compression and decompression subsystem 265, and further includes additional audio/visual input processors 380 and RGB analog to digital converters 390, to accommodate the additional video channels. As mentioned above with

reference to FIG. 13, for each channel (or RF carrier), channels a through n, there is a corresponding RF demodulator 275, namely, RF demodulators 275_a through 275_n. Referring to FIG. 17, the baseband input video signals from each of the RF demodulators 275_a through 275_n, such as NTSC/PAL composite video signals, are fed or input into corresponding audio/visual input processors 380, namely, audio/visual input processor 380_a through audio/visual input processor 380_n. As discussed above with reference to FIG. 6, the corresponding audio/visual input processors 380 convert the baseband input video signal, for each channel, into input analog RGB signals. The input analog RGB signals, for each channel a through n, are then converted into input digital RGB signals by corresponding RGB analog to digital converters 390, namely, RGB analog to digital converter 390_a through RGB analog to digital converter 390_n, and transferred to the video processing DSP 365. The video processing DSP 365 combines or multiplexes the digital RGB signals from each channel, creating one combined video signal, which (with the audio from the user audio interface 255) is then compressed and transferred to the microprocessor subsystem 260 and other components discussed above for protocol encoding and/or modulation, for subsequent delivery to the network 140. For example, in the preferred embodiment, each channel typically contains video information having a common interface format (CIF) or a quarter common interface format (QCIF); depending upon the chosen resolution, also for example, four such channels may be combined into one channel having four display windows utilizing a quarter common interface format (QCIF), before compression and transmission to the network 140.

Not illustrated in FIGs. 16 and 17, the video access apparatus 850 may also be expanded to provide for multiple output video channels, transmitting a plurality of radio

frequency output video signals to a plurality of videophones via the second communication channel 227. In that event, the video access apparatus 850 illustrated in FIG. 16 would be expanded to include a plurality of modulators 270, such as a modulator 270_a for video transmission on channel "a", a modulator 270_b for video transmission on channel "b", and so on. Similarly, the expanded audio/video compression and decompression subsystem 760 would be further expanded to include, correspondingly, a plurality of encoders 395 and a plurality of RGB digital to analog converters 370, again corresponding to each output channel. With such additional expansion, the audio/video conferencing system 705 may support multiple and independent video conference sessions. In this multiple independent video mode, the input video signals would also be processed independently, rather than being combined into a single combined video signal. For example, videophone 700_1 may communicate with remote location one and receive a video call on channel "a" while sending video on channel "b", while videophone 700_n may communicate with remote location two and independently receive a video call on channel "c" while independently sending video on channel "d". To support such multiple, independent video modes (also referred to as video bridging), the video access apparatus 850 may also include additional processing capability within the expanded audio/video compression and decompression subsystem 760, such as including a multipoint control unit (MCU) to provide processing support for such multiple independent video modes.

FIG. 18 is a flow diagram illustrating the video multiplexing method in accordance with the present invention. As discussed above, this methodology concerns the video signal to be protocol encoded and transmitted over the first communication channel 103 to the network 140, and

may be viewed as a special case or more narrow application of the method steps 510, 520, 530 and 540 discussed above with reference to FIG. 11. In addition, these multiplexing, combining, encoding and transmission steps (discussed below), like steps 510, 520, 530 and 540, also may proceed concurrently with and independently of the video receiving (from the first communication channel 103 and the network 140), protocol decoding, RF modulating and transmitting steps 505, 515, 525 and 535 (discussed above with reference to FIG. 11).

Referring to FIG. 18, the method begins, start step 900, with receiving an input audio signal, step 905, and with receiving a plurality of radio frequency input video signals, each carried on one of the plurality of channels a through n, namely, receiving a first radio frequency input video signal on a first channel, step 910_1, receiving a second radio frequency input video signal on a second channel, step 910_2, and so on, through receiving a n^{th} radio frequency input video signal on a n^{th} channel, step 910_n. Next, each of the plurality of radio frequency input video signals is demodulated to form a plurality of baseband input video signals. More particularly, the first radio frequency input video signal is demodulated to form a first baseband input video signal, step 915_1, the second radio frequency input video signal is demodulated to form a second baseband input video signal, step 915_2, and so on, through the n^{th} radio frequency input video signal being demodulated to form a n^{th} baseband input video signal, step 915_n. As indicated above, in the preferred embodiment, step 905 occurs in the user/audio interface 255, with the receiving steps 910_1 through 910_n and demodulating steps 915_1 through 915_n occurring in the plurality of RF demodulators corresponding to each channel a through n, namely, RF demodulators 275_a through 275_n. Next, each of the plurality of baseband input

of the present invention is designed to be compatible for use with other existing video conferencing systems, for example, those which may utilize either ISDN or POTS networks, rather than being solely compatible with one or the other (but not both. Moreover, the system of the present invention is user friendly, easy to install and use, and should be relatively less expensive for in-home purchase and use by consumers.

Another interesting feature of the apparatus and system embodiments of the present invention is the multiple functionality of the user interface, for example, the dual use of a telephone (as a user interface) for control of the video conference call and for the audio portion of the video conference call. This feature is also in stark contrast to the prior art systems, which typically require special switching and special network operations for call placement and call control. Such duality is in addition to the concomitant use of the telephone for POTS service. Yet another significant feature of the preferred embodiment of the present invention is the transparency of telephony operation, such that a user need not be aware of the video conferencing capability to place or receive a telephone call.

Other special features of the preferred embodiment include dual network and premise powering of the video access apparatus, or complete network powering, enabling continued functionality even during power outages. Yet another significant feature of the present invention is the "loop back" operation, such that the same system may also be utilized for surveillance, such as baby monitoring, in addition to conferencing. With the multiplexing capability of the present invention, the video from multiple cameras may be looped back, for example, to provide simultaneous surveillance of multiple locations. Another significant feature of the present invention is the independence of the audio portion from the video portion of an audio/video

conference. Moreover, the video conferencing capability illustrated is also protocol independent, such that a variety of communication protocols may be utilized and downloaded without user intervention.

Lastly, the multiplexing capability of the various videophone apparatuses 700 and 800 is truly unique. In addition to providing the functionality discussed above, such as videoconferencing functionality from multiple locations, the various videophone apparatuses 700 and 800, in conjunction with video access apparatus 850, also provides for multiplexing and combining the video signals from multiple locations into one combined video signal which may be transmitted over a network or looped back into the user premises.

From the foregoing, it will be observed that numerous variations and modifications may be effected without departing from the spirit and scope of the novel concept of the invention. It is to be understood that no limitation with respect to the specific methods and apparatus illustrated herein is intended or should be inferred. It is, of course, intended to cover by the appended claims all such modifications as fall within the scope of the claims.

What is claimed is:

1. A videophone apparatus, the videophone apparatus coupleable to a video access apparatus via a first communication channel for video reception and transmission, the videophone apparatus further coupleable to the video access apparatus via a second communication channel for audio reception and transmission, the videophone apparatus comprising:

a video monitor coupleable to the first communication channel;

a camera interface coupleable to the first communication channel;

a video camera coupled to the camera interface; and

a telephony module coupleable to the second communication channel.

2. A video access apparatus comprising:

a wireline network interface coupleable to a first communication channel for reception of a first protocol signal to form a received protocol signal and for transmission of a second protocol signal to form a transmitted protocol signal;

a radio frequency modulator to convert a baseband output video signal to a radio frequency output video signal;

a plurality of demodulators to convert a plurality of radio frequency input video signals to a plurality of baseband input video signals;

a user interface for reception of a first control signal of a plurality of control signals; and

a processor arrangement, the processor arrangement coupled to the wireline network interface, to the radio frequency modulator, to the plurality of radio frequency demodulators, and to the user interface, the processor arrangement responsive, through a set of program instructions, and in response to the first control signal, to convert the received protocol signal to the baseband output video signal and to an output audio signal, the processor arrangement further responsive to convert the plurality of baseband input video signals and an input audio signal to the second protocol signal.

3. An audio and video conferencing system, the audio and video conferencing system coupleable to a communication channel for audio and video transmission and reception, the audio and video conferencing system comprising:

a videophone apparatus; and

a video access apparatus coupled to the videophone apparatus.

4. A method of multiplexing video signals for transmission in audio/video conferencing, the method comprising:

(a) receiving an input audio signal;

(b) receiving a plurality of radio frequency input video signals, each corresponding to a channel;

(c) demodulating the plurality of radio frequency input video signals to correspondingly form a plurality of baseband input video signals;

(d) converting the plurality of baseband input video signals and the input audio signal to a second protocol signal; and

(e) transmitting the second protocol signal to form a transmitted protocol signal.

5. A video access apparatus comprising:

a wireline network interface coupleable to a first communication channel, the wireline network interface having an ISDN interface for reception of a first ISDN encoded H.32x digital audio/video protocol signal, for transmission of a second ISDN encoded H.32x digital audio/video protocol signal, and for transmission and reception of a digital telephony signal;

a radio frequency modulator to convert a baseband NTSC/PAL encoded composite output video signal to a radio

frequency amplitude modulated vestigial sideband output video signal;

a plurality of radio frequency demodulators to convert a plurality of radio frequency amplitude modulated vestigial sideband input video signals correspondingly to a plurality of baseband NTSC/PAL encoded composite input video signals;

an audio/user interface for reception of a first control signal of a plurality of control signals, for reception of an input analog audio signal and conversion of the input analog audio signal to an input digital audio signal, and for conversion of an output digital audio signal to an output analog audio signal and for output of the output analog audio signal;

a microprocessor subsystem coupled to the wireline network interface and the audio/user interface, the microprocessor subsystem responsive, through a set of program instructions, and in response to the first control signal, to decode the first ISDN encoded H.32x digital audio/video protocol signal to form an H.32x encoded output audio/video digital data stream, the microprocessor subsystem further responsive to convert a plurality of H.32x encoded input digital audio signals and an H.32x encoded input video signal data stream to the second ISDN encoded H.32x digital audio/video protocol signal; and

an audio/video compression and decompression subsystem coupled to the microprocessor subsystem, and further coupled to the radio frequency modulator and to the plurality of radio frequency demodulators, the audio/video compression and decompression subsystem responsive, through a set of program instructions, to convert and compress the plurality of baseband NTSC/PAL encoded composite input video signals to the plurality of H.32x encoded input video signal data stream, to convert and compress the input digital audio signal to the H.32x encoded

input digital audio signal, and to decompress and convert the H.32x encoded output audio/video digital data stream to the baseband NTSC/PAL encoded composite output video signal and the output digital audio signal.

6. A video access apparatus comprising:

a wireline network interface coupleable to a first communication channel, the wireline network interface having a telephony interface for reception and demodulation of a first V.x encoded H.32x analog audio/video protocol signal to form a first V.x encoded H.32x digital audio/video protocol signal, for modulation and transmission of a second V.x encoded H.32x digital audio/video protocol signal to form a transmitted V.x encoded H.32x analog audio/video protocol signal, and for transmission and reception of an analog telephony signal;

a radio frequency modulator to convert a baseband NTSC/PAL encoded composite output video signal to a radio frequency amplitude modulated vestigial sideband output video signal;

a plurality of radio frequency demodulators to convert a plurality of radio frequency amplitude modulated vestigial sideband input video signals correspondingly to a plurality of baseband NTSC/PAL encoded composite input video signals;

an audio/user interface for reception of a first control signal of a plurality of control signals, for reception of an input analog audio signal and conversion of the input analog audio signal to an input digital audio signal, and for conversion of an output digital audio signal to an output analog audio signal and for output of the output analog audio signal;

a microprocessor subsystem coupled to the wireline network interface and the audio/user interface; and

an expanded audio/video compression and decompression subsystem coupled to the microprocessor subsystem, and further coupled to the radio frequency modulator and to the plurality of radio frequency demodulators, the audio/video compression and decompression subsystem responsive, in conjunction with the microprocessor subsystem and the user audio interface, through a set of program instructions and in response to the first control signal, to decode, decompress and convert the first V.x encoded H.32x digital audio/video protocol signal to the baseband NTSC/PAL encoded composite output video signal and the output digital audio signal, and further responsive to encode, compress and convert the plurality of baseband NTSC/PAL encoded composite input video signals and the digital input audio signal to the second V.x encoded H.32x digital audio/video protocol signal.

The Patent Office

ŁS

Application No: GB 9720970.4
Claims searched: 1

Examiner: Ken Long
Date of search: 29 December 1997

Patents Act 1977
Search Report under Section 17

Databases searched:

UK Patent Office collections, including GB, EP, WO & US patent specifications, in:

UK Cl (Ed.O): H4K (KOD3, KF56 & KF56A)

Int Cl (Ed.6): H04M (3/56 & 11/08)
H04N (7/14 & 7/15)

Other: NONE

Documents considered to be relevant:

Category	Identity of document and relevant passage		Relevant to claims
X	GB 2294377 A	AMULET (page 1 lines 3-6, page 3 lines 20-23 and page 4 line 22 to page 5 line 20)	1
X	GB 2280822 A	MATSUSHITA (page 11 line 22 to page 13 line 7)	1
X	WO 95/12281 A1	THOMSON-CSF (see components 1-17 Fig 1)	1
X	US 5541640	LARSON (column 8 lines 43-63)	1
X	US 5524110	INTEL (column 4 line 43 to column 5 line 4 and column 5 lines 48-56)	1
X	US 5473366	CANON (column 13 lines 6-53)	1

X	Document indicating lack of novelty or inventive step	A	Document indicating technological background and/or state of the art.
Y	Document indicating lack of inventive step if combined with one or more other documents of same category.	P	Document published on or after the declared priority date but before the filing date of this invention.
		E	Patent document published on or after, but with priority date earlier than, the filing date of this application.
&	Member of the same patent family		